

STIC-Biotech/ChemLib

137546

mej

From: Whiteman, Brian
Sent: Wednesday, November 10, 2004 10:02 AM
To: STIC-Biotech/ChemLib
Subject: sequence search

10/069,386, Alaoui-Jamali et al.
2/19/02

search SEQ ID NO: 2 against databases.

Thank you,

Brian Whiteman
Remsen, 2D14
mail box 2C18
Patent Examiner - Art Unit 1635
United States Patent and Trademark Office
(571) 272-0764

RECEIVED
NOV 10 2004
STIC-BIOTECH/ChemLib
(STIC)

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA Sequence: # _____
AA Sequence :# _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2004, 13:49:05; Search time 156 Seconds
(without alignments)
450.711 Million cell updates/sec

Title: US-10-069-386-2
Perfect score: 1033
Sequence: 1 MEGSLKRXHSDLEEEERWE.....APGSWENELDHIMEIILGS 196

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 202273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A Geneseq 23Sep04:*
- 1: geneseqp1980s:*
 - 2: geneseqp1980s:*
 - 3: geneseqp2000s:*
 - 4: geneseqp2001s:*
 - 5: geneseqp2002s:*
 - 6: geneseqp2003as:*
 - 7: geneseqp2003bs:*
 - 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1033	100.0	196	4	AAB35402
2	1026	99.3	196	4	AAM93922
3	1026	99.3	196	8	ADL32052
4	791	76.6	157	4	AAG74742
5	645.5	62.5	142	8	ABO60075
6	402	38.9	111	8	ABO60074
7	216	20.9	236	2	AAY36004
8	216	20.9	236	8	ADP19312
9	213	20.6	236	3	AAY44362
10	213	20.6	236	4	AAM93724
11	213	20.6	236	4	AAG89292
12	213	20.6	236	8	ADL31644
13	213	20.6	236	4	AAM35550
14	211	20.4	236	2	AAY31829
15	205	19.8	222	8	ABO60367
16	205	19.8	236	2	AAY02619
17	136	13.2	237	5	AAO22897
18	125	12.1	314	5	ABP65076
19	106	10.3	241	4	AM39725
20	106	10.3	254	3	AB58258
21	106	10.3	254	4	AAG73682
22	106	10.3	254	4	AM41511
23	105.5	10.2	578	4	AM38707
24	105.5	10.2	620	7	ADN95361
25	105.5	10.2	650	4	AM38706

Ade28201 Human MDD
Aam40492 Human pol
Aam40493 Human pol
Adn95128 Human IEC
Abb06375 Human CHD
Aao22898 Human hae
Abbi1713 Human KIA
Abp69285 Human pol
Abg22868 Novel hum
Add46841 Human pro
Ade56294 Human pro
Adj66562 P13 Kinas
Adb50606 Human pro
Abr53281 Protein s
Adk63588 Disease t
Aam78754 Human pro
Adp22490 Sea-squir
Adj32004 Human T-C
Aam79738 Human pro
Adh75872 Corn C-re

ALIGNMENTS

RESULT 1
AAB35402
ID AAB35402 standard; protein; 196 AA.
AC AAB35402;
XX
XX 23-MAY-2001 (first entry)
DE Replication protein A binding transcriptional activator 1 RBT1.
XX RBT1; replication protein A binding transcriptional activator 1; RPA32;
KW gene therapy; apoptosis; cancer; leukaemia.
XX
XX Unidentified.
XX WO200114546-A2.
XX
XX 01-MAR-2001.
XX
XX 17-AUG-2000; 200OWO-CA000948.
XX
XX 19-AUG-1999; 99US-0149472P.
XX (TRAN-) CENT TRANSLATIONAL RES IN CANCER.
PA
XX
PI Alaoui-Jamali MA, Cho JM;
XX
DR WPI; 2001-218447/22.
XX N-PSDB; AAF28052.
XX
PT Novel replication protein A binding transcriptional activator 1 gene,
PT useful for treating neoplastic disorders such as cancer and in gene
PT therapy.
XX
PS Disclosure; Fig 1; 16pp; English.
XX
XX The present invention provides the protein and coding sequences of the
XX replication protein A binding transcriptional activator 1 (RBT1). The
XX protein is capable of inducing apoptosis. The sequences are useful in the
XX gene therapy and other methods of treatment of cancer, including
XX leukaemias. The present sequence is the RBT1 protein
XX
SQ Sequence 196 AA;
Query Match 100.0%; Score 1033; DB 4; Length 196;
Best Local Similarity 100.0%; Pred. No. 1.5e-92;
Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEGGLKRKHSLEEEERWSPAGLQSYQQALLRISLDKVQSLGPRAPSLRRHVLHN 60
 DB 1 MEGGLKRKHSLEEEERWSPAGLQSYQQALLRISLDKVQSLGPRAPSLRRHVLHN 60
 QY 61 TLQQLQALRLAPALPPEPLFLGEEDFSLSATIGSILRELDTSMDGTPPQNPVTPLG 120
 DB 61 TLQQLQALRLAPALPPEPLFLGEEDFSLSATIGSILRELDTSMDGTPPQNPVTPLG 120
 QY 121 LQNEVPPODPVFLEALSSRYLGSGLDDFFLDIDTSAVEKEPARAPPEPHNLFCAPGS 180
 DB 121 LQNEVPPODPVFLEALSSRYLGSGLDDFFLDIDTSAVEKEPARAPPEPHNLFCAPGS 180
 QY 181 WENNELDHIMEIILGS 196
 DB 181 WENNELDHIMEIILGS 196

RESULT 2

AM93922
 ID AM93922 standard; protein; 196 AA.

XX AM93922;

DT 06-NOV-2001 (first entry)

DE Human polypeptide, SEQ ID NO: 4085.

XX Human; full length cDNA; cDNA synthesis; oligo-capping.

XX Homo sapiens.

XX EPI130094-A2.

PD 05-SEP-2001.

XX 07-JUL-2000; 2000EP-00114089.

XX 08-JUL-1999; 99JP-00194486.

XX 11-JAN-2000; 2000JP-00118774.

XX 02-MAY-2000; 2000JP-00183765.

XX (HELI-) HELIX RES INST.

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI; 2001-524255/58.

XX N-PSDB; AAK94884.

XX 830 Primers useful for synthesizing full length cDNA clones and their use

XX in genetic manipulation.

XX Claim 8; SEQ ID NO 4085; 1380pp + Sequence Listing; English.

XX The invention relates to primers for synthesizing full length cDNA

XX clones. 830 cDNA molecules encoding a human protein have been isolated

XX and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have

XX been determined. Primers for synthesizing the full length cDNA are useful

XX for clarifying the function of the protein encoded by the cDNA. The full

XX length clones were obtained by construction of full length enriched cDNA

XX libraries that were synthesised by the oligo-capping method. The primers

XX enable the production of the full length cDNA easily without any special

XX methods. The present sequence is a polypeptide encoded by a full length

XX human cDNA of the invention. Note: The sequence data for this patent did

XX not form part of the printed specification, but was obtained in CD-ROM

XX format directly from EPO

XX Sequence 196 AA;

XX Query Match 99.3%; Score 1026; DB 4; Length 196;

XX Best Local Similarity 99.5%; Pred. No. 7.5e-92;

XX Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEGGLKRKHSLEEEERWSPAGLQSYQQALLRISLDKVQSLGPRAPSLRRHVLHN 60
 DB 1 MEGGLKRKHSLEEEERWSPAGLQSYQQALLRISLDKVQSLGPRAPSLRRHVLHN 60
 QY 61 TLQQLQALRLAPALPPEPLFLGEEDFSLSATIGSILRELDTSMDGTPPQNPVTPLG 120
 DB 61 TLQQLQALRLAPALPPEPLFLGEEDFSLSATIGSILRELDTSMDGTPPQNPVTPLG 120
 QY 121 LQNEVPPODPVFLEALSSRYLGSGLDDFFLDIDTSAVEKEPARAPPEPHNLFCAPGS 180
 DB 121 LQNEVPPODPVFLEALSSRYLGSGLDDFFLDIDTSAVEKEPARAPPEPHNLFCAPGS 180
 QY 181 WENNELDHIMEIILGS 196
 DB 181 WENNELDHIMEIILGS 196

RESULT 3

ADL32052

ID ADL32052 standard; protein; 196 AA.

XX ADL32052;

XX 20-MAY-2004 (first entry)

DE Human protein encoded by a full length cDNA clone SeqID 4085.

XX human; medicine; signal transduction; glycoprotein; transcription;

XX oligo-capping method.

XX Homo sapiens.

XX EPI396543-A2.

XX 10-MAR-2004.

XX 07-JUL-2000; 2003EP-00025638.

XX 08-JUL-1999; 99JP-00194486.

XX 11-JAN-2000; 2000JP-00118774.

XX 02-MAY-2000; 2000JP-00183865.

XX 07-JUL-2000; 2000EP-00114089.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI; 2004-204755/20.

XX N-PSDB; ADL32051.

XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full

XX length human cDNAs.

XX Example 1; SEQ ID NO 4085; 1340pp; English.

XX This invention relates to a novel primers useful for synthesising full

XX length cDNA molecules that encode human proteins. Specifically, it refers

XX to secretory or membrane proteins that are potential therapeutic agents/

XX target molecules in the field of medicine, and in particular genes

XX encoding proteins that are associated with signal transduction,

XX glycoproteins and transcription. The present invention describes a method

XX for efficiently cloning a full length human cDNA from both the 5' and 3'

XX ends using the oligo-capping method. This polypeptide sequence is a full

XX length human protein of the invention.

XX Sequence 196 AA;

XX Query Match 99.3%; Score 1026; DB 8; Length 196;

XX Best Local Similarity 99.5%; Pred. No. 7.5e-92;

XX Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX 1 MEGGLKRKHSLEEEERWSPAGLQSYQQALLRISLDKVQSLGPRAPSLRRHVLHN 60

Db 1 MVGGLKRRHSDLEESERWENSPAGLSQYQOALLRISLDKQVRSLSGPRAPSLRRHVLHN 60
Qy 61 TLQOQOALRLAPALPEPLFLGEEFSLSATIGSLRELDTSMTGTPPQNPVTPLG 120
Db 61 TLQOQOALRLAPALPEPLFLGEEFSLSATIGSLRELDTSMTGTPPQNPVTPLG 120
Qy 121 LQNEVPPODPVPFLEALSRYLGDGLDDFFLDIDTSAVEKEPARAPPEPHNLFCAPGS 180
Db 121 LQNEVPPODPVPFLEALSRYLGDGLDDFFLDIDTSAVEKEPARAPPEPHNLFCAPGS 180
Qy 181 WEWNELDHIMEIILGS 196
Db 181 WEWNELDHIMEIILGS 196

RESULT 4

AAG74742
ID AAG74742 standard; protein; 157 AA.

XX AAG74742;
AC AAG74742;

XX 03-SEP-2001 (first entry)

XX Human colon cancer antigen protein SEQ ID NO:5506.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX Colorectal carcinoma.

XX Homo sapiens.

XX WO200122920-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US026524.

XX 29-SEP-1999; 99US-0157137P.

XX 03-NOV-1999; 99US-0163280P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI; 2001-235357/24.

XX N-PSDB; AAH341147.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers.

XX Claim 11; Page 7098-7099; 9803pp; English.

XX AAH22943 to AAH37195 and AAG73514 to AAG77788 represent human colon
XX cancer-associated nucleic acid molecules (N) and proteins (P), where the
XX proteins are collectively known as colon cancer antigens. The colon
XX cancer antigens have cytostatic activity and can be used in gene therapy
XX and vaccine production. N and P may be used in the prevention, diagnosis
XX and treatment of diseases associated with inappropriate P expression. For
XX example, N and P may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of P by expressing inactive proteins or to
XX supplement the patient's own production of P. Additionally, N may be used
XX to produce the colon cancer-associated Ps, by inserting the nucleic acids
XX into a host cell and culturing the cell to express the proteins. N and P
XX can be used in the prevention, diagnosis and treatment of colorectal
XX carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent
XX sequences used in the exemplification of the present invention. N.B.
XX Pages 666 to 682 and page 7053 of the sequence listing were missing at
XX time of publication, meaning no sequences are present for SEQ ID NO:1027
XX to 1052, 7921 and 7922

XX Sequence 157 AA;

Query Match 76.6%; Score 791; DB 4; Length 157;

Best Local Similarity 98.0%; Pred. No. 5.1e-69;

Matches 150; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 44 SLGPPAPSLRRHVLHNTLQOQOALRLAPALPEPLFLGEEFSLSATIGSLRELD 103

Db 5 ALGPPAPSLRRHVLHNTLQOQOALRLAPALPEPLFLGEEFSLSATIGSLRELD 64

Qy 104 TSMGTEPPQNPVTPFLGONEVPPQDPVPFLEALSRYLGDGLDDFFLDIDTSAVEKEP 163

Db 65 TSMGTEPPQNPVTPFLGONEVPPQDPVPFLEALSRYLGDGLDDFFLDIDTSAVEKEP 124

Qy 164 ARAPPEPHNLFCAPGSWEWNELDHIMEIILGS 196

Db 125 ARAPPEPHNLFCAPGSWEWNELDHIMEIILGS 157

RESULT 5

ABO60075

ID ABO60075 standard; protein; 142 AA.

XX ABO60075;

XX 29-JUL-2004 (first entry)

XX Human genome derived single exon protein #6309.

XX Human; gene expression; single exon probe; microarray;
XX alternative splicing event; genomic alteration.

XX Homo sapiens.

XX US2003194704-A1.
XX 16-OCT-2003.

XX 03-APR-2002; 2002US-00029386.

XX 03-APR-2002; 2002US-00029386.

XX (PENN/) PENN S G.

XX (RANK/) RANK D R.

XX (HANZ/) HANZEL D K.

XX Penn SG, Rank DR, Hanzel DK;

XX WPI; 2004-119264/12.

XX New human genome-derived single exon nucleic acid probes useful for human
XX gene expression analysis, for identifying or characterizing alternative
XX splicing events, for assessing genomic alterations or as tools for
XX surveying tissues.

XX Claim 45; SEQ ID NO 33709; 80pp; English.

XX The invention relates to a nucleic acid probe for measuring human gene
XX expression, comprising any of the 27,400 fully defined nucleotide
XX sequences in the specification, or their complements or fragments, and
XX encoding at least 8 amino acids of any of the 688 amino acid sequences
XX fully defined in the specification. The probe is a single exon probe that
XX hybridizes under high stringency conditions to a nucleic acid molecule
XX expressed in human cells or tissues. Also included are a spatially-
XX addressable set of single exon nucleic acid probes for measuring human
XX gene expression (comprising a plurality of single exon nucleic acid
XX probes cited above, where each of the plurality of probes is separately
XX and addressably isolatable or amplifiable from the plurality), a single
XX exon microarray for measuring human gene expression, a method of
XX measuring human gene expression, a vector comprising the single exon
XX probe cited above, an ORF-encoded peptide comprising at least 8
XX contiguous amino acids of any of the above-mentioned amino acid
XX sequences (optionally with conservative amino acid substitutions), an
XX isolated antibody that binds specifically to a peptide cited above,

XX methods of selling and/or licensing single exon probes or microarrays to

PD	24-JUN-1999.	XX	10-JUN-2004.	XX
PF	17-DEC-1998; 98WO-IB002122.	PF	15-OCT-2001; 2001US-00978360.	PF
XX		XX		XX
PR	17-DEC-1997; 97US-0069957P.	PR	17-DEC-1998; 98WO-IB002122.	PR
PR	09-FEB-1998; 98US-0074121P.	PR	09-FEB-1999; 99WO-IB000282.	PR
PR	13-APR-1998; 98US-0081563P.	PR	21-JUN-2000; 2000WO-IB000951.	PR
PR	10-AUG-1998; 98US-0096116P.	PR	15-SEP-2000; 2000US-00663600.	PR
XX		XX		XX
PA	(GEST) GENSET.	PA	(GEST) GENSET SA.	PA
XX		XX		XX
XX	Bougueleret L, Duclert A, Dumas Milne Edwards J;	XX	Dumas Milne Edwards J, Bougueleret L, Jobert S, Clusel C;	XX
PI		PI	Duclert A;	PI
XX		XX		XX
DR	WPI; 1999-385906/32.	DR	WPI; 2004-440404/41.	DR
DR	N-PSDB; AAY97688.	DR	N-PSDB; ADP18907.	DR
XX		XX		XX
PT	New isolated human secreted proteins.	PT	New isolated polynucleotide encoding secreted polypeptide, useful for	PT
XX		XX	gene therapy, or in diagnostic procedures to identify individuals having	XX
PS	Claim 9; Page 330-331; 516pp; English.	PS	genetic diseases resulting from abnormal expression of the genes.	PS
XX		XX		XX
CC	This sequence is encoded by an extended human secreted protein coding	CC	Claim 2; SEQ ID NO 568; 113pp; English.	CC
CC	sequence of the invention. The secreted proteins can be used in treating	CC	The invention relates to human cDNA sequences that encode human secreted	CC
CC	or controlling a variety of human conditions. The secreted proteins may	CC	proteins. The invention also relates to an antibody that specifically	CC
CC	act as cytokines or may affect cellular proliferation or differentiation	CC	binds to a polypeptide of the invention and a method of binding the	CC
CC	or may act as immune system regulators, haematopoiesis regulators, tissue	CC	polypeptide to an antibody. The polynucleotides are useful for expressing	CC
CC	growth regulators, regulators of reproductive hormones or cell movement	CC	the entire secreted proteins which they encode and for distinguishing	CC
CC	or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or	CC	human tissues and cells from non-human tissues and cells, and for	CC
CC	tumour inhibition activity. The DNAs can be used in forensic procedures	CC	distinguishing between human tissues and cells that do or do not express	CC
CC	to identify individuals or in diagnostic procedures to identify	CC	the polynucleotides comprising the cDNAs. The polynucleotides and	CC
CC	individuals having genetic diseases resulting from abnormal expression of	CC	polypeptides are useful in forensic procedures or diagnostic procedures	CC
CC	the genes corresponding to the extended cDNAs. They are also useful for	CC	to identify individuals with genetic diseases resulting from abnormal	CC
CC	constructing a high resolution map of the human chromosomes. They can	CC	expression of the genes corresponding to the cDNAs. The sequences are	CC
CC	also be used for gene therapy to control or treat genetic diseases	CC	also useful in gene therapy to control or treat genetic diseases. This	CC
XX		CC	sequence represents a human secreted polypeptide of the invention. Note:	CC
SQ	Sequence 236 AA;	CC	The sequence data for this patent did not form part of the printed	CC
	Query Match 20.9%; Score 216; DB 2; Length 236;	CC	specification but was obtained in electronic format from USPTO at	CC
	Best Local Similarity 32.6%; Pred. No. 1.3e-12;	XX	seqdata.uspto.gov/sequence.html.	XX
	Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;			
QY	4 GLKRKSHDLREE-----ERNEWSPAGLOSQQA-----LLRISLDKVORSGLGRAPS 51	QY	4 GLKRKSHDLREE-----ERNEWSPAGLOSQQA-----LLRISLDKVORSGLGRAPS 51	
Db	5 GLKRKREEEKEEPLAVDSW-WLDPGHAAVAQAPPAVASSLFDLSVLKHLHSLQXSPD 63	Db	5 GLKRKREEEKEEPLAVDSW-WLDPGHAAVAQAPPAVASSLFDLSVLKHLHSLQXSPD 63	
QY	52 LRHHVLIHNTLQQLAALRLAPALPPEPL-----FLGEEDFSLSATIGSILRE 101	QY	52 LRHHVLIHNTLQQLAALRLAPALPPEPL-----FLGEEDFSLSATIGSILRE 101	
Db	64 LRHLVXNLTIRIQAS--MAPAAALPPVPTPPAAPXVADNLLASSDAALSASMAXLLED 121	Db	64 LRHLVXNLTIRIQAS--MAPAAALPPVPTPPAAPXVADNLLASSDAALSASMAXLLED 121	
QY	102 LDTSMGTEPPQNVTPLGLONEVPPQD----PVFLEAL-----SSRYLGDSDLDDFFL 152	QY	102 LDTSMGTEPPQNVTPLGLONEVPPQD----PVFLEAL-----SSRYLGDSDLDDFFL 152	
Db	122 L-SHIEGLSQAPQ-----LADEGPPGRSIGGXPPXPLGALDGLGATGCLDNGLEGLFE 175	Db	122 L-SHIEGLSQAPQ-----LADEGPPGRSIGGXPPXPLGALDGLGATGCLDNGLEGLFE 175	
QY	153 DIDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEIILGS 196	QY	153 DIDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEIILGS 196	
Db	176 DIDTSMYDNLWAPASGLKPGPED---GPGKEAPELDEALDYLMDVLVGT 225	Db	176 DIDTSMYDNLWAPASGLKPGPED---GPGKEAPELDEALDYLMDVLVGT 225	
RESULT 8		RESULT 9		
ADP19312		AAY44362		
ID	ADP19312 standard; protein; 236 AA.	ID	AAY44362 standard; protein; 236 AA.	
XX		XX		
AC	ADP19312;	AC	AAY44362;	
XX		XX		
DT	26-AUG-2004 (first entry)	DT	14-MAR-2000 (first entry)	
XX		XX		
DE	Human secreted polypeptide #163.	DE	Human cell cycle regulation protein-3.	
XX		XX		
KW	Human; secreted protein; genetic disease.	KW		
XX		XX		
OS	Homo sapiens.	OS		
XX		XX		
PN	US2004110939-A1.	PN		
XX		XX		

CECRP-3; cell cycle regulation protein-3; cell proliferation;
 cell proliferative disease; cancer; atherosclerosis; cirrhosis;
 hepatitis; psoriasis; immune system disorder; allergy; asthma;
 acquired immune deficiency syndrome; Crohn's disease; Blast method;
 Rheumatoid arthritis; gene therapy; chromosomal mapping.

XX Homo sapiens.

Key Location/Qualifiers
 Modified-site 44
 /note= "Potential phosphorylation site"
 Modified-site 60
 /note= "Potential phosphorylation site"
 Modified-site 73
 /note= "Potential phosphorylation site"
 Modified-site 98
 /note= "Potential phosphorylation site"
 Modified-site 117
 /note= "Potential phosphorylation site"
 Modified-site 123
 /note= "Potential phosphorylation site"
 Modified-site 180
 /note= "Potential phosphorylation site"

XX WO9964593-A2.

XX 16-DEC-1999.

XX 08-JUN-1999; 99WO-US012906.

XX 08-JUN-1998; 98US-0088695P.

XX (INCY-) INCYTE PHARM INC.

XX Bandman O, Lal P, Tang YT, Corley NC, Guegler KJ, Baughn MR;
 Patterson C;

XX WPI: 2000-105887/09.

XX N-PSDB; AAZ29482.

XX Novel regulatory proteins, for diagnosis, treatment and prevention of
 cell proliferative and immune system diseases.

XX Claim 1; Page 70; 88pp; English.

XX The present sequence is cell cycle regulation protein-3 (CECRP-3). Prints
 analytical method was used to identify this protein. CECRPs are
 activators of cell proliferation or inhibitors of cellular processes that
 modulate proliferation. They are used to treat or prevent cell
 proliferative diseases like cancers, atherosclerosis, cirrhosis,
 hepatitis, psoriasis, immune system disorders (e.g. acquired immune
 deficiency syndrome, allergy, asthma, Crohn's disease, rheumatoid
 arthritis). Antibodies are raised to screen for specific binding agents.
 The corresponding nucleic acid is used in gene therapy, chromosomal
 mapping and isolation of related sequences

XX Sequence 236 AA;

Query Match 20.6%; Score 213; DB 3; Length 236;
 Best Local Similarity 32.6%; Pred. No. 2.6e-12;
 Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY 4 GLKRKSHDLEEE-----ERWESPAGLQSYQQA-----LLRLSLDKVQSLGFRAPS 51
 DB 5 GLKRKREEEKEPLAVDSW-WLDPGHAAVQAAPPVASSSLFDLSVLKHLHSLQSQSPD 63
 QY 52 LRRLVLIHNTLQQLAQLRLAPALPPEPL-----FLGEDFSLSATIGSLIRE 101
 DB 64 LRHLVLVNTLRLRIQAS--MAPAAALPPVPSPAPSVADNLTASSDAALSASMASLLED 121
 QY 102 LDTSMGTEPPQNPVTFPLGQNEVPP-----QDPVFLEAL--SSRYLGDSGLDDFFL 152
 DB 122 L-SHIEGLSCAQQP-----LADEGPPRSIGGAAPSLGALDGLGATGCLLDGGLGLE 175

QY 153 DDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMFIILGS 196
 DB 176 DDTSMYDNELWAPASEGLKPGPED---GPKKEAPELDEALDYLDVLTGT 225

RESULT 10

AM93724
 ID AAM93724 standard; protein; 236 AA.

XX AC AAM93724;

XX DT 06-NOV-2001 (first entry)

XX DE Human polypeptide, SEQ ID NO: 3677.

XX KW Human; full length cDNA; cDNA synthesis; oligo-capping.

XX OS Homo sapiens.

XX PN EPI130094-A2.

XX PD 05-SEP-2001.

XX PF 07-JUL-2000; 2000EP-00114089.

XX PR 08-JUL-1999; 99JP-00194486.

XX PR 11-JAN-2000; 2000JP-00118774.

XX PR 02-MAY-2000; 2000JP-00183765.

XX PA (HELI-) HELIX RES INST.

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI: 2001-524255/58.

XX N-PSDB; AAK94674.

XX 830 Primers useful for synthesizing full length cDNA clones and their use

XX in genetic manipulation.

XX Claim 8; SEQ ID NO 3677; 1380pp + Sequence Listing; English.

XX The invention relates to primers for synthesizing full length cDNA
 clones. 830 cDNA molecules encoding a human protein have been isolated
 and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 been determined. Primers for synthesizing the full length cDNA are useful
 for clarifying the function of the protein encoded by the cDNA. The full
 length clones were obtained by construction of full length enriched cDNA
 libraries that were synthesised by the oligo-capping method. The primers
 enable the production of the full length cDNA easily without any special
 methods. The present sequence is a polypeptide encoded by a full length
 human cDNA of the invention. Note: The sequence data for this patent did
 not form part of the printed specification, but was obtained in CD-ROM
 format directly from EPO

XX Sequence 236 AA;

Query Match 20.6%; Score 213; DB 4; Length 236;
 Best Local Similarity 32.6%; Pred. No. 2.6e-12;
 Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY 4 GLKRKSHDLEEE-----ERWESPAGLQSYQQA-----LLRLSLDKVQSLGFRAPS 51
 DB 5 GLKRKREEEKEPLAVDSW-WLDPGHAAVQAAPPVASSSLFDLSVLKHLHSLQSQSPD 63
 QY 52 LRRLVLIHNTLQQLAQLRLAPALPPEPL-----FLGEDFSLSATIGSLIRE 101
 DB 64 LRHLVLVNTLRLRIQAS--MAPAAALPPVPSPAPSVADNLTASSDAALSASMASLLED 121
 QY 102 LDTSMGTEPPQNPVTFPLGQNEVPP-----QDPVFLEAL--SSRYLGDSGLDDFFL 152
 DB 122 L-SHIEGLSCAQQP-----LADEGPPRSIGGAAPSLGALDGLGATGCLLDGGLGLE 175

64 LRHLVNVTLRRIQAS--MAPAAALPPVPSPAAPSAVDNLLASSDAALSASWASLLED 121
 102 LDTSMGTEPPQNVPVPLGLONEVP-----QDPVFLEAL--SSRYLGDGSLDDFFL 152
 122 L-SHIEGLSQAPQP-----LADEGPPGRSISGGAAPSLGALDGLGPAFGCLLDGLGLELFE 175
 153 DDTSAVEKE---PARAPPEP-PHNLFCAPGSW-----WNELDHIMEIILGS 196
 176 DDTSMYDNELWAPASEGLKGPED---GPKKEAPELDEALDYLMVDVLVGT 225

RESULT 13

AAW25550
 ID AAW25550 standard; protein; 278 AA.

XX AC AAW25550;

DT 16-OCT-2001 (first entry)

DE Human protein sequence SEQ ID NO:1065.

XX Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
 XX anti-inflamatory; anti-rheumatic; anti-arthritis; immunosuppressive;
 XX anti-bacterial; endocrine; cardiant; central nervous system; virucide;
 XX anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
 XX antiagregant; haemostatic; vulnerary; antiulcer; osteopathic; cykostatik;
 XX dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
 XX neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
 XX immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
 XX antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
 XX cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
 XX genetic disease; haematopoietic disorder; platelet disorder; asthma;
 XX thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
 XX allergic rhinitis; diabetes; multiple sclerosis; depression;
 XX Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 XX neurological disorder.

OS Homo sapiens.

XX W0200153455-A2.

XX 26-JUL-2001.

XX 22-DEC-2000; 2000WO-US035017.

XX 23-DEC-1999; 99US-00471275.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-00552317.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-457603/49.

XX N-PSDB; AAW99491.

XX Isolated human polynucleotides encoding polypeptides, useful for the
 XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection.

XX Claim 20; Page 214; 1217pp; English.

XX AAH99166 to AAH99904 encode the human proteins given in AAW25225 to
 CC AAW25963. The proteins can have activities based on the tissues and cells
 CC they are expressed in, such as: anti-inflamatory; anti-rheumatic;
 CC anti-arthritis; immunosuppressive; antibacterial; endocrine; cardiant;
 CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
 CC cardiovascular; antianaemic; antiagregant; haemostatic; vulnerary;
 CC antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis

CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders

XX Sequence 278 AA;

Query Match 20.6%; Score 213; DB 4; Length 278;

Best Local Similarity 32.6%; Pred. No. 3.2e-12;

Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY 4 GLKRKHSLEEEE-----ERWESFAGLSQYQA-----LLRISLDKVSRLGPRAPS 51

DB 47 GLKRKEEKEKEPLAVDSW-WLDPGHAAVAQAPPAVASSSLFDLSVLKHLHSLQSEPD 105

QY 52 LRRHVLHNTLQQLQAALRLAPAPALPPEPL-----FLGEEDFSLSATIGSLRE 101

DB 106 LRHLVNVTLRRIQAS--MAPAAALPPVPSPAAPSAVDNLLASSDAALSASWASLLED 163

QY 102 LDTSMGTEPPQNVPVPLGLONEVP-----QDPVFLEAL--SSRYLGDGSLDDFFL 152

DB 164 L-SHIEGLSQAPQP-----LADEGPPGRSISGGAAPSLGALDGLGPAFGCLLDGLGLELFE 217

QY 153 DDTSAVEKE---PARAPPEP-PHNLFCAPGSW-----WNELDHIMEIILGS 196

DB 218 DDTSMYDNELWAPASEGLKGPED---GPKKEAPELDEALDYLMVDVLVGT 267

RESULT 14

AAV31829

ID AAV31829 standard; protein; 236 AA.

XX AC AAV31829;

XX 06-DEC-1999 (first entry)

XX Human adult blood secreted protein g21_1.

XX Secreted protein; g21_1; human; therapy; diagnosis; vaccine; blood.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Misc-difference 1

XX Peptide 19..21

XX Protein 32..236

XX Domain 40

XX Peptide 67..79

XX Protein 80..236

XX Domain 80

XX Misc-difference 137

XX Domain 150

XX WO9947555-A1.

XX 23-SEP-1999.

XX 18-MAR-1999; 99WO-US005939.
 XX 20-MAR-1998; 98US-00078803P.
 PR 17-MAR-1999; 99US-00078803.
 XX (GEM) GENETICS INST INC.
 PA Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Agostino MJ, Steininger RJ;
 XX WPI: 1999-562059/47.
 DR N-PSDB; AA219894.
 XX New polynucleotides derived from murine fetal cell cDNA libraries,
 PT potentially used as, e.g. vaccines.
 XX Claim 13(a); Page 94; 107pp; English.
 XX This is the predicted amino acid sequence of a novel human secreted
 CC protein, g21.1, as deduced from an isolated adult blood cDNA clone (see
 CC AA219894). The invention provides new human secreted proteins (see
 CC AA219894-38) and polynucleotides (see AA219893-901) isolated from foetal
 CC cell, adult blood, and foetal kidney cDNA libraries. They are
 CC predicted to have biological activities which would make them suitable
 CC for treating, preventing or ameliorating medical conditions in humans and
 CC animals, although no supporting data are given. Suggested activities
 CC include nutritional, cytokine, tissue growth, cell proliferation and
 CC differentiation, immunostimulant (e.g. as vaccine), immunosuppressive,
 CC haematopoiesis regulating, activin or inhibin, chemotactic or
 CC chemokinetic, haemostatic or thrombolytic, receptor/ligand activity,
 CC antiinflammatory, cadherin or tumour invasion suppressor, and tumour
 CC inhibition activities
 XX Sequence 236 AA;
 SQ
 Query Match 20.4%; Score 211; DB 2; Length 236;
 Best Local Similarity 32.6%; Pred. No. 4.1e-12;
 Matches 76; Conservative 35; Mismatches 70; Indels 52; Gaps 13;
 QY 4 GLKXKSDLEEE-----ERWENSPAGLSQYQQA-----LLRSLDKVQSLGPRAPS 51
 DB 5 GLKXKSDLEEE-----ERWENSPAGLSQYQQA-----LLRSLDKVQSLGPRAPS 51
 QY 52 LRRHVLHNTLQLOQAALRLAPALPPEPL-----FIGEEDFSLSATIGSLRE 101
 DB 64 LRLHLVNVNLTTRIQAS--MAPAAALPPVPSPAPSPVADNLLASSDAALSASMASLLED 121
 QY 102 LDTSMGTEPPQNPVTPLQNEVPP-----QPDVFLEAL--SSRYLGDSDLDDPFL 152
 DB 122 L-SHIEGLSQAPQ-----LAXEGPPGRSIGGAAPSLGALDGLGATGCLDDGLEGLFE 175
 QY 153 DITSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEILGS 196
 DB 176 DITSMYDNLWAPASEGLKPGED---GPGKEAPELDEALDYLMDVLVGT 225
 RESULT 15
 ABO60367
 ID ABO60367 standard; protein; 222 AA.
 XX ABO60367;
 XX 29-JUL-2004 (first entry)
 XX Human genome derived single exon protein #601.
 XX Human; gene expression; single exon probe; microarray;
 XX alternative splicing event; genomic alteration.
 XX Homo sapiens.
 XX US2003194704-A1.
 XX
 XX 16-OCT-2003.
 XX 03-APR-2002; 2002US-00029386.
 XX 03-APR-2002; 2002US-00029386.
 XX (PENN/) PENN S G.
 XX (RANK/) RANK D R.
 XX (HANZ/) HANZEL D K.
 XX Penn SG, Rank DR, Hanzel DK;
 XX WPI: 2004-119264/12.
 XX New human genome-derived single exon nucleic acid probes useful for human
 PT gene expression analysis, for identifying or characterizing alternative
 PT splicing events, for assessing genomic alterations or as tools for
 PT surveying tissues.
 XX Claim 45; SEQ ID NO 34001; 80pp; English.
 XX The invention relates to a nucleic acid probe for measuring human gene
 CC expression, comprising any of the 27,400 fully defined nucleotide
 CC sequences in the specification, or their complements or fragments, and
 CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
 CC fully defined in the specification. The probe is a nucleic acid molecule
 CC hybridises under high stringency conditions to a nucleic acid molecule
 CC expressed in human cells or tissues. Also included are a spatially-
 CC addressable set of single exon nucleic acid probes for measuring human
 CC gene expression (comprising a plurality of single exon nucleic acid
 CC probes cited above, where each of the plurality of probes is separately
 CC and addressably isolatable or amplifiable from the plurality), a single
 CC exon microarray for measuring human gene expression, a method of
 CC measuring human gene expression, a vector comprising the single exon
 CC probe cited above, an ORF-encoded peptide comprising at least 8
 CC contiguous amino acids of any of the above-mentioned amino acid
 CC sequences (optionally with conservative amino acid substitutions), an
 CC isolated antibody that binds specifically to a peptide cited above,
 CC methods of selling and/or licensing single exon probes or microarrays to
 CC a customer desiring to measure gene expression, and a computer-readable
 CC human gene expression data by subscription, and a computer-readable
 CC storage medium which contains a database having a plurality of records
 CC (each record including data on the expression of a single exon probe
 CC cited above). The probe, methods and apparatus are useful in gene
 CC expression analysis. The probes may be used as tools for surveying
 CC tissues to detect the presence of expressed messages that contain their
 CC specific exon, or in constructing genome-derived single exon microarrays.
 CC In addition, the probes are used in identifying and characterising
 CC alternative splicing events, in detecting and characterising gross
 CC alterations in the genomic locus that includes their exon, in assessing
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,
 CC or in expressing the ORF-encoded peptide. The present sequence is a human
 CC single exon probe protein of the invention. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?docID=20030194704
 XX Sequence 222 AA;
 SQ
 Query Match 19.8%; Score 205; DB 8; Length 222;
 Best Local Similarity 32.6%; Pred. No. 1.5e-11;
 Matches 75; Conservative 34; Mismatches 69; Indels 52; Gaps 13;
 QY 4 GLKXKSDLEEE-----ERWENSPAGLSQYQQA-----LLRSLDKVQSLGPRAPS 51
 DB 5 GLKXKSDLEEE-----ERWENSPAGLSQYQQA-----LLRSLDKVQSLGPRAPS 51
 QY 52 LRRHVLHNTLQLOQAALRLAPALPPEPL-----FIGEEDFSLSATIGSLRE 101
 DB 64 LRLHLVNVNLTTRIQAS--MAPAAALPPVPSPAPSPVADNLLASSDAALSASMASLLED 121
 QY 102 LDTSMGTEPPQNPVTPLQNEVPP-----QPDVFLEAL--SSRYLGDSDLDDPFL 152

us-10-069-386-2.rag

Mon Nov 15 17:33:06 2004

```

122 L-SHIEGUSQAPQ-----LADEGPPGRSIGGAAPSLGALDILGPATGCLDDGLEGLFE 175
153 DIDISAVEKE---PARAPPEP-PHNLFCAPGSWE-----KNELDHIMEII 193
176 DIDTSMYDNEIWAFASEGLKXGPEP---GFGKEEAPBELDEALDYLMVDL 222

```

Search completed: November 15, 2004, 14:03:14
Job time : 159 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: November 15, 2004, 13:49:50 ; Search time 198 Seconds
(without alignments)
569.563 Million cell updates/sec

Title: US-10-069-386-2
Perfect score: 1033
Sequence: 1 MEGGLKRXHSDLEEEERWE.....APGSWENELDHIMEILGS 196

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1033	100.0	196	2 Q9UW9	Q9UW9 homo sapien
2	1026	99.3	196	2 Q96CQ2	Q96CQ2 homo sapien
3	871	84.3	197	2 Q9ERC3	Q9ERC3 mus musculu
4	217	21.0	236	1 STD1 MOUSE	Q9J110 mus musculu
5	213	20.6	236	1 STD1 HUMAN	Q9UHV2 homo sapien
6	208	20.1	236	2 Q6P71	Q6P771 rattus norv
7	208	20.1	236	2 AAH61808	AAH61808 rattus no
8	166	16.1	244	2 Q9DC22	Q9DC22 mus musculu
9	140.5	13.6	309	1 STD2 MOUSE	Q9JJ95 mus musculu
10	136.5	13.2	248	2 Q6GWB1	Q6GM81 xenopus lae
11	136	13.2	237	1 CCA4 MOUSE	Q9CWM2 mus musculu
12	136	13.2	237	2 AAH55824	AAH55824 mus muscu
13	133	12.9	361	2 Q6NVD9	Q6NVD9 brachydanio
14	133	12.9	361	2 AAH67135	AAH67135 brachydan
15	133	12.9	383	2 Q7ZZ27	Q7ZZ27 brachydanio
16	125	11.1	314	1 STD2 HUMAN	Q14140 homo sapien
17	115	11.1	1194	2 Q7GM68	Q7GM68 rattus norv
18	115	11.1	1194	2 RAD14305	RAD14305 rattus no
19	113.5	11.0	911	2 Q80TJ8	Q80TJ8 mus musculu
20	112.5	10.9	246	2 Q6DFP1	Q6DFP1 xenopus lae
21	105.5	10.2	620	2 Q9BQ18	Q9BQ18 homo sapien
22	105.5	10.2	760	2 Q76N32	Q76N32 homo sapien
23	105.5	10.2	760	2 BAA25508	BAA25508 homo sapi
24	105	10.2	241	1 CCA4 HUMAN	Q9BX18 homo sapien
25	104.5	10.1	740	2 Q9UP22	Q9UP22 homo sapien
26	101.5	9.8	728	1 P85B HUMAN	O00459 homo sapien
27	101.5	9.8	728	2 AAH70082	AAH70082 homo sapi
28	100	9.7	846	2 Q6NV99	Q6NV99 brachydanio
29	100	9.7	846	2 AAH68186	AAH68186 brachydan
30	98.5	9.5	314	2 Q6NXL2	Q6NXL2 mus musculu
31	98.5	9.5	314	2 AAH67019	AAH67019 mus muscu

ALIGNMENTS

RESULT 1

Q9UW9 PRELIMINARY; PRT; 196 AA.
AC Q9UW9
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE RPA-binding trans-activator.
GN Name=RBT1;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20440390; PubMed=10982866;
RA Cho J.M., Song D.J., Alsouli-Jamali M.A.;
RT "RBT1, a novel transcriptional co-activator, binds the second subunit
of Replication Protein A";
RL Nucleic Acids Res. 28:3478-3485 (2000).
DR EMBL; AF192529; AAF05761.1; -.
DR InterPro; IPR009263; SERTA.
DR Pfam; PF06031; SERTA; 1.
SQ SEQUENCE 196 AA, 21798 MW; 33433EF7F8A9EBA7 CRC64;

Query Match 100.0%; Score 1033; DB 2; Length 196;

Best Local Similarity 100.0%; Pred. No. 1.le-72;

Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEGGLKRXHSDLEEEERWEKSPAGLQSYQQALLRISLDKQVRSILGPRAPSLRRHVLHN 60
DB 1 MEGGLKRXHSDLEEEERWEKSPAGLQSYQQALLRISLDKQVRSILGPRAPSLRRHVLHN 60
QY 61 TLQQLQAALRLAPAPALPEPLFLGEEEDFSLSATIGSLRLDTSMDGTPEPPQNPVPLG 120
DB 61 TLQQLQAALRLAPAPALPEPLFLGEEEDFSLSATIGSLRLDTSMDGTPEPPQNPVPLG 120
QY 121 LQNEVPPQDPVFLSALSRVYLGSDGLDDFFLDITSAVEKEPARAPPEPHNLFCAFGS 180
DB 121 LQNEVPPQDPVFLSALSRVYLGSDGLDDFFLDITSAVEKEPARAPPEPHNLFCAFGS 180
QY 181 WENNELDHIMEILGS 196
DB 181 WENNELDHIMEILGS 196

RESULT 2

Q96CQ2 PRELIMINARY; PRT; 196 AA.
AC Q96CQ2
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-OCT-2004 (Tremblrel. 28, Last annotation update)
DE RPA-binding trans-activator.

AC Q9JL10; Q925E6; Q9D888;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE SERTA domain-containing protein 1 (Transcriptional regulator
DE interacting with the PHD-bromodomain 1) (TRIP-Bri) (CDK4-binding
DE protein p34SEI1) (SEI-1)
GN Name=Seitad; Synonyms=Seil;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RP Ohtani N., Hara E.;
RA "Cloning of mouse SEI-1 cDNA.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A., FUNCTION, AND INTERACTIONS.
RX MEDLINE=21231173; PubMed=1131592;
RA Heu S.-I., Yang C.M., Sim K.G., Hentschel D.M., O'Leary E.,
RA Bonventre J.V.;
RA "TRIP-Bri: a novel family of PHD zinc finger- and bromodomain-
RA interacting proteins that regulate the transcriptional activity of
RT E2F-1/DP-1.";
RL EMBO J. 20:2273-2285(2001).
[3]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo, and Small intestine;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito K., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gajobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriber L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Brack D., Brucic V., Chocho A.C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Glessi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Walstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang L., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
[4]
RN SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein W.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.K., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalski U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Acts at E2F-responsive promoters to integrate signals
CC provided by PHD- and/or bromodomain-containing transcription
CC factors. Stimulates E2F-1/DP-1 transcriptional activity. Renders
CC the activity of cyclin D1/CDK4 resistant to the inhibitory effects
CC of p16(INK4a).
CC -!- SUBUNIT: Interacts with the PHD-bromodomain of TIF1, TRIM28/TIF1B
CC and p300/CBP. Binds to Dp1. Also interacts with CDK4.
CC -!- SIMILARITY: Belongs to the TRIP-Bri family.
CC -!- SIMILARITY: Contains 1 SERTA domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF218291; AAF27653.1;
CC EMBL: AF366400; AAK52829.1;
CC EMBL: AF366401; AAK52830.1;
CC EMBL: AK004022; BAB3130.1;
CC EMBL: AK008303; BAB25588.1;
CC EMBL: BC016077; AAH16077.1;
CC MGD: MGI:1913438; Sertad1.
CC InterPro: IPR009263; SERTA.
CC Pfam: PF06031; SERTA.1.
CC Transcription regulation.
CC DOMAIN 45 82
CC CONFLICT 9 10
CC CONFLICT 12 12 E -> K (in Ref. 2; AAK52829).
CC SEQUENCE 236 AA; 25136 MW; 989ADF8299DE84C5 CRC64;
Query Match 21.0%; Score 21.7; DB 1; Length 236;
Best Local Similarity 31.5%; Pred. No. Se-09; Indels 68; Gaps 12;
Matches 76; Conservative 33; Mismatches 64;
QY 4 GLKRRHSDLEEEERWE-----W-----SPAGLQS-----YQALLRLSLDKVQSLGPR 48
Db 5 GLKRR-----REEETMEALSVDSDLPSPHVAQTPPTVASSSLFSLVVKLHSLRQS 60
QY 49 APSLRHVLHNTLQOQALRLAPALPEPL-----FLGEEDFSLSATIGSI 98
Db 61 EPDLRLHLVNVNLRRIQASME--PAPVLPPEPIQPPAPSVADSLASDAGLSMASL 118
QY 99 LRELDTSDMGTEPPQNPVTPGLQNEVPP-----QDPVVFLEAL--SSRYLGSGLDD 149
Db 119 LEDLNLHIEDLNCAPQQA-----DEGPFGRSIGISNLGALDLGATGCLLDGLEG 172
QY 150 FFLDIDTSAVEKE-----PAR-----APPEPHNLFCAPGSEWNEIDHINEILG 195
Db 173 LFEDIDTSMYDSELWLPASEGLKPGFNGPAKEPP-----ELDEALDYLDMDVLVG 224
QY 196 S 196
Db 225 T 225
RESULT 5
ID STD1 HUMAN STANDARD; PRT; 236 AA.
AC Q9UHV2; Q9BUE7;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last sequence update)

```

entities requires a license agreement (See http://www.isb-sib.ch/announce/)
or send an email to licens@isb-sib.ch).

-----
CC EMBL; AF117959; AAF08349.1; -
CC CC EMBL; AF366402; AAK52831.1; -
CC DR DR EMBL; AY130860; AAM77800.1; -
CC DR EMBL; BC002670; AAH02670.1; -
CC DR Gensw; HGNC:17932; SKRTAD1.
CC DR GO; GO:000828; P:positive regulation of cell proliferation; TAS.
CC DR GO; GO:0000379; P:regulation of CDK activity; TAS.
CC DR InterPro; IPR009263; SERTA.
CC DR Pfam; PF06031; SERTA; 1.
CC KW Polymorphism; Transcription regulation.
CC FT DOMAIN 45 82 SERTA.
CC FT VARIANT 31 31 A -> T (in dbSNP:268697) .
CC FT /FTID=VAR_015881.
CC SQ SEQUENCE 236 AA; 24673 MW; 92F2EA328F15B1 CRC64;

Query Match 20.6%; Score 213; DB 1; Length 236;
Best Local Similarity 32.6%; Pred No. 1e-08;
Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY 4 GLKRKHSLEBEE---ERWENSPAGLSYQQA-----LLRISLDKVQSLSGRAPS 51
Db ||||| : || : || : || : || : || : || : || : || : || : ||
5 GLKRKEEEEEKEPLAVDSW-WLDPCHAAVAQPAVAVASSSLFDLSVLKLHSLQQSEPD 63
QY 52 LRHHVLIHTLQQLCAALRLAPALPEPL-----FLGEEDFSLSATIGSILRE 101
Db ||||| : || : || : || : || : || : || : || : || : || : ||
64 LRHLVLVNTLRRIQAS--MAPAAALPPVPSPAAPSVDNULNASDAALSMSASLLED 121
QY 102 LDTSMGDTGPONPVTPVTLGLQNEVP-----QPDPVFLEAL--SSRYLGDSGLDDFFL 152
Db ||||| : || : || : || : || : || : || : || : || : || : ||
122 L-SHIEGLSQAPQ-----LADEGPPGRSIGGAAPSLGALDDELGPATGCLDDGLEGLFE 175
QY 133 DIDSITSAVEKE---PARAPPETPHNLFCAPSW-----WNELDHIMEILLGS 196
Db ||||| : || : || : || : || : || : || : || : || : || : ||
176 DIDTSMYNELWAPASEGLKPQPED---GPGKEEAPELDSEALDYMDVLVST 225

RESULT 6
Q6P771 PRELIMINARY; PRG; 236 AA.
AC Q6P771
ID Q6P771
IC Q6P771
AD Q6P771
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
OX [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RP MEDLINE=23288257; PubMed=12477932;
RX Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Strausberg R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusica K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyik S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Griinwood J., Smith J., Myers R.M., Butterfield Y.S.,
RA Krzywicki M.I., Skalska U., Smutitz J., Scherch A., Schein G.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human

```



```
RT and mouse cDNA sequences." ;
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP TISSUE=Prostate;
RC SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC061808; AAH61808.1; -.
RW Hypothetical protein.
SQ SEQUENCE 236 AA; 25419 MW; B8B7502E0473ED8B CRC64;

Query Match 20.1%; Score 208; DB 2; Length 236;
Best Local Similarity 31.9%; Pred. No. 2.5e-08;
Matches 76; Conservative 33; Mismatches 67; Indels 62; Gaps 13;

QY 4 GLKXKSHDLSEERWE-----W-----SPAGLSYQOALLRLISLKVQSLG 46
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5 GLKXK-----REEETMESLSVDSWMLDQSRPAVAQTPTVAS--SSLFSLVVKLHSLR 58
QY 47 PRAPSLRRHVLHNTLQOQAALRLAPALPPEPL-----FIGEEDFSLSATIG 96
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 59 QSEPDRLHLVLTNTLRIQASME--PTTVLPPEPTOPTAPSVADNLFSSSDAGLSASMA 116
QY 97 SILRELDTSMDGTEPPQNPVTPGLQNEVPP-----QDPVFLEAL--SSRYLGDGSL 147
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 117 SLLEDLSHIEDLNQVPOQA-----DEGPPGRSVGGVLPNLGALDGLLGPATGCLLDGGL 170
QY 148 DDFELDIDTSVAEKE---PARAPPEP-PHNLFCAPGSWEW-----NELDHIMEILGS 196
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 171 EGLFEDIDTSWYDSELWLPASEGLKVPEN---GPAKEETPELDEALDYLMVLTG 225

RESULT 7
AAH61808 PRELIMINARY; PRT; 236 AA.
ID AAH61808;
AC AAH61808;
DT 02-MAR-2004 (TREMREL. 27, Created)
DT 02-MAR-2004 (TREMREL. 27, Last sequence update)
DT 02-MAR-2004 (TREMREL. 27, Last annotation update)
DE Hypothetical protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gharatne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzyzanski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences." ;
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
```

```

RD SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=20530913; PubMed=11076861;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Kono H., Akiyama J., Nishi K., Kitsuana T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayashizaki Y.;
RA "RKEN integrated sequence analysis (RISA) system-384-format
RA sequencing pipeline with 384 multicapillary sequencer.";
RA Genome Res. 10:1757-1771 (2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Haragaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shingagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toyota T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RA Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RA EMBL: AK002324; BAE22013.1; -
RA MGD: MGI:1913438; Sertad1.
RA InterPro: IPR009263; SERTA.
RA Pfam: PF06031; SERTA; 1.
SQ SEQUENCE 244 AA; 26068 MW; 0F464F6419D3362A6 CRC64;
Query Match 16.1%; Score 166; DB 2; Length 244;
Best Local Similarity 25.9%; Pred. No. 4.8e-05;
Matches 67; Conservative 33; Mismatches 55; Indels 104; Gaps 13;
OY 4 GLRKHSDLEERERWE-----W-----SPAGLQS-----YQALLRISDKVCRSLGPR 48
DB 5 GLKEN-----GEEDTMAISVDSCLDPSHPAQAQPTTVASSLSFDLSVVKLHSLRQS 60
OY 49 APSLRRLVHTLTLQQLQAALRLAPALPPEPL-----FLGEEDFSLSATIGSI 98
DB 61 EPDLRLHLVAVNTLRLRQAQME--PAPVLPPEPTQPAPSPVADSLASSDAGLSASWASL 118
OY 99 LRELDTSMDGTEPPQ-----NPVTP-----LG-----LQNEV 125
DB 119 LEDLNHIEDLNQAPQADSGPPGRSGTGGISPNLGAIDLQQLAVCWTDWACRLTSI 178
OY 126 PP-----QPDPVFLEALSSRYLGLSDGLDFDLDTDSVAVEKEPARAPPPEPHNLFCA 177
DB 179 PCCITVNVGYQP-----LRVSSAAPENGPAK--EEPP-----208
OY 178 PGSEWNELDHMEIILGS 196
DB 209 --ELDEAELTDLMDVLVGT 225
RESULT 9
STD2_MOUSE STANDARD; Q91WL3; Q92S55; PRT; 309 AA.
AC Q9JUG5; Q8C609; Q91WL3; Q92S55;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE SERTA domain-containing protein 2 (Transcriptional regulator
DE interacting with the PHD-bronco domain 2) (TRIP-Br2).
GN Name=Sertad2; Synonyms=Kiaa0127;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RA Osada N., Kusuda J., Tanuma R., Ito A., Hirata M., Sugano S.,
RA Hashimoto K.;
RA "Isolation of full-length cDNA clones from mouse brain cDNA library
RA made by oligo-capping method.";
RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yananaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisels K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Guscinich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Kongaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Perte G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sadelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyszewski-Boris A., Yanagisawa M., Yang L., Yang L.,
RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RA "Analysis of the mouse transcriptome based on functional annotation of
RA 60,770 full-length cDNAs.";
RA Nature 420:563-573 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Boeck S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.M.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RA and mouse cDNA sequences.";
RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [4]
RP SEQUENCE OF 129-309 FROM N.A.
RX MEDLINE=21231173; PubMed=11331592;
RA Hsu S.-I., Yang C.M., Sim X.G., Hentschel D.M., O'Leary E.,
RA Bonventre J.V.;
RA "TRIP-Br: a novel family of PHD zinc finger- and bromodomain-
RA interacting proteins that regulate the transcriptional activity of
RA E2F-1/DP-1.";

```

```

RL EMBO J. 20:2273-2285(2001).
CC -I- FUNCTION: Acts at E2F-responsive promoters to integrate signals
CC provided by pRb- and/or bromodomain-containing transcription
CC factors (By similarity)
CC -I- SIMILARITY: Belongs to the TRIP-BR family.
CC -I- SIMILARITY: Contains 1 SERTA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB041541; BAA95026.1; -
DR EMBL; AK076787; BAC36480.1; -
DR EMBL; BC014726; AH14726.1; -
DR EMBL; AF366403; AAK52832.1; -
DR PIR; PTO566; PTO566.
DR MGD; MGI:1931026; Sertad2.
DR InterPro; IPR009263; SERTA.
DR Pfam; PF06031; SERTA; 1.
KW Transcription regulation.
FT CONFLICT 138 140 CTL -> FTF (in Ref. 4).
FT CONFLICT 142 142 A -> S (in Ref. 3).
FT CONFLICT 145 145 P -> S (in Ref. 4).
FT CONFLICT 234 234 T -> A (in Ref. 1).
SQ SEQUENCE 309 AA; 33312 MW; D4178688F0DF8F00 CRC64;

Query Match 13.6%; Score 140.5; DB 1; Length 309;
Best Local Similarity 24.8%; Pred. No. 0.0063;
Matches 77; Conservative 29; Mismatches 85; Indels 119; Gaps 13;

QY 2 EGGKXKHSLEEEERWENSPAGLOS-----YQALLRISLDKQVRSGLGPRAPSLRRH 55
Db 4 KGG-KKKDEHEDGLEGKIVSPDGRSVYTLQRTQTFINSLMKLYNHRPLTEPSLOKT 62
QY 56 VLIHNTLQLOALRL-----APALPPEPLFGE----- 86
Db 63 VLINMLRRIQELKQEGSLRPAFTPSQPSNLSDSYQEAAPPA--PHPCDLGSGTPTLE 120
QY 87 -----EDFSLSATIGSILRELDTSMDG 108
Db 121 ACLTPASLBEDNDFTCTLQAVHPAATFLSSAALPAEKDSFSSALDEI-EELCPTSTS 179
QY 109 TEPPQN--PVTPGLQNEVPPO-----PDVPFLEAL-----SSRYLGDGSLD 148
Db 180 TEAHTAAPGPKGTSESSVQKPEEGRTDSDRFMSLPGNFBITTSQGLTDLTLD 239
QY 149 D-PFLDIDTSAVEKEPA-----RAPPE-----PPHNLFCAPG---SWEWNEL 186
Db 240 DILFADIDTSMYDFDPTCSASGTASKWAPVSADLLKTLAPYSNQVPASQPFKMDLTLE 299
QY 187 DHIMEILGS 196
Db 300 DHIMEVLGS 309

RESULT 10
Q6GM81 PRELIMINARY; PRT; 248 AA.
AC Q6GM81;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
CX NCBI_TaxID=8355;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Posak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Heiton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A.C., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative."
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Klein S., Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC074195; AAH74195.1; -
DR InterPro; IPR009263; SERTA.
DR Pfam; PF06031; SERTA; 1.
KW Hypothetical protein.
SQ SEQUENCE 248 AA; 27319 MW; E703D7E9105DF08C CRC64;

Query Match 13.2%; Score 136.5; DB 2; Length 248;
Best Local Similarity 30.9%; Pred. No. 0.0098;
Matches 58; Conservative 25; Mismatches 58; Indels 47; Gaps 8;

QY 17 ERWENSPAGLOSQYQALLRISLDKQVRSGLGPRAPSLRRHVLHNTLQLOALRL---AP 73
Db 10 ERECCSIPAIOQH--CLMNI SLVKLHRSRHEVPDLRFVLA VANTLRRLQNLQVEOCAP 67
QY 74 -----APALPPEPLFGEEDFSLSATIGSILRELDTSMDGTEPPQ 113
Db 68 DMWKTSECTRSALVVPESKXPALNTEDPLSSMDASLYSSITILEDLN-NFEGLS--S 124
QY 114 NPVTPLGLQNEVPPODPVFLEA-----LSSRY-----LGDGSLDFFLDID 155
Db 125 SPLQIEDDQLCAPKANPVGSAEWMVKLASSSLSSPYLLGGLNLDN-LEDIFEDID 183
QY 156 TSAVEKEP 163
Db 184 TSMYSDSP 191

RESULT 11
CCA4_MOUSE
ID CCA4_MOUSE STANDARD; PRT; 237 AA.
AC Q9CWM2; Q921E8; Q99MP6;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Cell division cycle associated protein 4 (Hematopoietic progenitor
DE protein).

```

GN Name=Cdca4; Synonyms=HEPP; Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 FC STRAIN=C57BL/6J; TISSUE=Bone marrow;
 FX MEDLINE=21375891; PubMed=11462882; DOI=10.1006/bcmd.2001.0434;
 RA Abdullah J.M., Jing X., Spassov D.S., Nachtman R.G., Jurecic R.;
 RT "Cloning and characterization of Hepp, a novel gene expressed preferentially in hematopoietic progenitors and mature blood cells";
 RL Blood Cells Mol. Dis. 27:667-676(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 FC STRAIN=C57BL/6J; TISSUE=Mesonephros;
 FX MEDLINE=22354683; PubMed=12456851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schrim L.M., Karapin A., Matsuda H., Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S., Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S., Gaasterland T., Gariboldi M., Glasi C., Godzik A., Gough J., Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D., Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L., Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A., Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H., Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S., Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M., Sandelin A., Schneider C., Sempie C.A., Setou M., Shimada K., Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C., Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L., Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I., Miyazaki A., Sakai K., Sakaki D., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Birney E., Hayashizaki Y.;
 RA "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs";
 RL Nature 420:563-573(2002).
 RN [3]
 RP SEQUENCE OF 13-237 FROM N.A.
 FC TISSUE=Breast tumor;
 FX STRAUSBERG R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Maman A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzyewski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: May be involved in molecular regulation of hematopoietic stem cells and progenitor cell lineage commitment and differentiation.

CC -!- TISSUE SPECIFICITY: Expressed preferentially in hematopoietic progenitors and mature blood cells. Expressed at low levels in the heart, lung, spleen, and thymus and at a higher level in muscle.
 CC -!- DEVELOPMENTAL STAGE: Developmentally regulated. Preferential expression in both fetal and adult hematopoietic progenitors and mature blood cells during embryonic and adult hematopoiesis.
 CC -!- SIMILARITY: Belongs to the TRIP-B family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; AK010535; BAB27012.2; -
 CC EMBL; AF322238; AAK31074.1; -
 CC EMBL; AK032980; BAC28109.1; -
 CC EMBL; EC012983; AAI12933.1; ALT_INIT.
 CC MGD; MGI:1919213; Cdc44.
 CC InterPro; IPR009263; SERTA.
 CC Pfam; PF06031; SERTA; 1.
 CC CONFLICT 196 196 S -> T (in Ref. 3).
 CC CONFLICT 213 213 T -> A (in Ref. 3).
 CC SEQUENCE 237 AA; 26107 MW; 52EC046BDFD326E CRC64;
 Query Match 13.2%; Score 136; DB 1; Length 237;
 Best Local Similarity 26.1%; Pred. No. 0.01;
 Matches 63; Conservative 32; Mismatches 86; Indels 60; Gaps 9;
 QY 4 GLKRSKDLSEEEERNEWSPA-GLQSYQQALLRLSLDKVQSLGPRAPSLRRHVLHNTL 62
 DB 5 GLKRYGQDESGVGFGTVPYSYLSQ--RQSLDKSLVKLQJLCHMLVFNLCRSVLIANTV 62
 QY 63 QQLQAARL-----LAP--APALPPPLFLGEEFDSLSATIGLSILRELDTSMDGTSPQ 113
 DB 63 RQIOEEMSQDGVWHGMAPQNVDRAPVRLVSTE---ILCRTVGAEEHFAPELEDAPLQ 119
 QY 114 NPVTPLGLQNEVPQPDP-----VFLEALSSRYLGDSDLDF 151
 DB 120 NSVSELPVGSAPQAPQPSLWNSDQFQNSGFSQKSLDOIFETLENK--NSSSVSELP 177
 QY 152 LDIDTSVAVEKEPAR-----APPEPPHNLFCAGSWENWELDHIMEII 193
 DB 178 SDVDSYVDLDTVLGTGWSGKSSLCNGLEGFAAATPPSPSTCKS---DLAELDHVVVEIL 234
 QY 194 L 194
 DB 235 V 235
 RESULT 12
 AAH55824
 ID AAH55824 PRELIMINARY; PRT; 237 AA.
 AC AAH55824;
 DT 02-MAR-2004 (TRENBLrel. 27, Created)
 DT 02-MAR-2004 (TRENBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TRENBLrel. 27, Last annotation update)
 DE Cell division cycle associated 4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=EYE;
 EX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Maman A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzyewski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: May be involved in molecular regulation of hematopoietic stem cells and progenitor cell lineage commitment and differentiation.

RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carinanci P., Prange C.,
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulláhy S.J.,
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.
RA	Fahey J., Helton E., Kettenan M., Madan A., Rodrigues S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.C., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA	Krzywinski M.I., Skalska U., Smalhus D.E., Schnerrch A., Schein J.E.,
RA	Jones S.J., Marra M.A.;
RT	"Generation and initial analysis of more than 15,000 full-length human
RT	and mouse cDNA sequences";
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Eye;
RA	Strausberg R.;
DR	Submitted (AUG-2003) to the ENBL/GenBank/DBJ databases.
RL	ENBL; BC055824; AAH55824.1; -.
KW	Gail division.
QY	SEQUENCE 237 AA; 26107 MW; 52EC046EFDDF326E CRC64;
Query Match 13.2%; Score 136; DB 2; Length 237;	
Best Local Similarity 26.1%; Pred. NO. 0.01;	
Matches 63; Conservative 32; Mismatches 86; Indels 60; Gaps 9	
QY	4 GLKRXHSLDEEEERKEWSPA-GLQSVOQALLRISLDKVORSIGPRAPSURRHVLIHTVL 62
Dd	5 GLKKRYGDOEEGVGGTVPYSYLQ--RQSLDMSLVKLQLCHMLVEPNLCRSVLIAITV 62
QY	63 QLOQAALR-----LAP--APALPEPLFLGEEDFSLSATIGSILRELTSDMGTEPPQ 113
Dd	63 RLQIQEMSODGWHGXAQNVDRAVERLNSTE---ILCRTVGAEEHPAPELEDAPIQ 119
QY	114 NPVTPLGLQNEVPPQDP-----VFLEAUSSRYLGDSGLDDFF 151
Dd	120 NSVSSELPVGSAGPGQRNPQSSLWMDSPOENRGFSQKSLDIQIFETLENK--NSSSVLELF 177
QY	152 LDIDSAVEKEPAR-----APPEPHNLFCAPGSWEWNELDHIMEII 193
Dd	178 SUVDSGYDLDTVLITGMMSGTKSSLCNGLEGFAATPPPSTCKS----DLAELDHVVEIL 234
QY	194 L 194
Dd	235 V 235
RESULT 13	
Q6XND9	ID Q6XND9 PRELIMINARY; PRT; 361 AA.
AC	Q6XND9;
DT	05-JUL-2004 (TrEMBLrel. 27, Created)
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE	Hypothetical protein sertad2.
GN	Name=sertad2;
CS	Brachydanio rerio (Zebrafish) (Danio rerio)..
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC	Cyprinidae; Danio.
OX	NCBI_TaxID=7955;
RN	[1]---
RP	SEQUENCE FROM N.A.
RC	TISSUE=kidney;
RC	MDLLINE=22388257; PubMed=12477932;
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA	Altschul S.F., Zeeberg B., Buotow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heih F.,
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

```

RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalks J., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
EL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR ENBL; BC067135; AAH67135.1;
DR InterPro; IPR009263; SERTA.
DR Pfam; PF06031; SERTA, 1.
KW Hypothetical protein.
SQ SEQUENCE: 361 AA; 37965 MW; 59CE364B098A43BF CRC64;

Query Match 12.9%; Score 133; DB 2; Length 361;
Best Local Similarity 22.3%; Pred. No. 0.029;
Matches 81; Conservative 35; Mismatches 71; Indels 176; Gaps 17;

QY 4 GLRKHSDLDEEBERWEW-----SPAGIQ-----SY-----QQALLRISLDKQRSUGPRAPS 51
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5 GAKRK-----LDEDEGLEGKALAAAGAGELSKVSYTLQRQTIFNMSLMKLYNHRVATPS 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 52 LRHVLHNTLOOLQALR-----PEPL-----FLGEDFSLSA-----L 71
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 62 LKRVLINNMLRIQELKQEGNLRPLFFPPPPDDPVDESFEQPAFSLVSMVAPPI 121
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 72 APALP-----PEPL-----FLGEDFSLSA----- 93
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 122 SQSPALSASLTPSSGLSNPAPLEACLTPAPLEEDNVSLCTSPSPPLAPPATSRISLPS 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 94 ----TIGSILRELD-----TSMGDTGEP-----QNPVTEL 119
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 182 VARDSPFSALDEIEELCPSLPTATSPAGTATSPLOLCPPLNSGALDSDCKSPK 241
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 120 --GL-----QNEVPPQPD--PVFLF-----ALSKRYLGDGSLD--FFLIDIDTSAVEK 161
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 242 LEGLVPLAERSAVENTPEPLPNSLDMSTSPSSSGFLTDLALDILFADIDTSYDF 301
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 162 EP-----ARAPPEPHNLFCAQPSWENNELDHIMEIL 193
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 302 DPCTSSGGAAPSKLAPMWTADLELKTFSYSGAAPVSSN---QPFKXDLTELDHIMEVL 358
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 194 LGS 196
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 359 VGS 361

RESULT 14
AAH67135 PRELIMINARY; PRT; 361 AA.
AC AAH67135;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein gi:2c10in13.8.
SI:ZC10IN13.8.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]

```

PP SEQUENCE FROM N.A.
 PG TISSUE=Kidney; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrincci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences";
 RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [2]
 PP SEQUENCE FROM N.A.
 PG TISSUE=Kidney;
 RA Strausberg R.;
 RA Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC067135; AAH67135.1; -
 RN Hypothetical protein.
 SQ SEQUENCE 361 AA; 37965 MW; 59CE364B098A43BF CRC64;
 Query Match 12.9%; Score 133; DB 2; Length 361;
 Best Local Similarity 22.3%; Pred. No. 0.029;
 Matches 81; Conservative 35; Mismatches 71; Indels 176; Gaps 17;
 QY 4 GLKRXHSLDEEEERWEW----SPAGLQ-----SY----QQALLRISLDKQVRSIGLGRAPS 51
 DB 5 GAKRK---LDEDEGLEKALAGAGAGLKVSYTLQRTIFNWSLMKLYNHRVTEPS 61
 QY 52 LRRHVLHNTLQQLQAALR-----PEPL-----FLGEEDFSLA-----L 71
 DB 62 LEKRVLLNNMLRRIQDELKQEGNLRPLFFPPPPDDPDVDSFREPQAPFVLSMWVAPPI 121
 QY 72 APAPALP-----TIGSILRELD-----TSMDGTTEP-----QNPVTPL 119
 DB 122 SQSPALSASSLTPSSSGLSNPALEACLTAPLLEEDNVSLCTSPSPPLAPPAPTSRLSPS 181
 QY 94 ---TIGSILRELD-----TSMDGTTEP-----QNPVTPL 119
 DB 182 VARDFSSALDEIEELCPSPPLTATSPAGTSPSPQLCPPSLNSGALDSKDCSKPCSPK 241
 QY 120 --GL-----QNEVPPQPD---PVFLE-----ALSSRYLGDSGLDD--FFLDIDTSVAVEK 161
 DB 242 LEGVLPLAERSAVPNTPETLPNLSLDMSTSPSASSSGFLTDALDDILFADIDTSMYDF 301
 QY 162 EP-----ARAPPEPPHNLFCAPGSWENWELDHIMEII 193
 DB 302 DPCTSSSGAAPSKLAPMVTADLLKTFSPYSGAAPAVSSN---QPFKMDLTDLHIMEVL 358
 QY 194 LGS 196
 DB 359 VGS 361

RESULT 15

Q7Z227 ID Q7Z227 PRELIMINARY; PRT; 383 AA.
 AC Q7Z227;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE S1:zC101N13.8 (Novel protein similar to human solute carrier family 1

DE (Neuronal/epithelial high affinity glutamate transporter, system Xag),
 DE member 1 (S1C1A1)).
 GN Name=S1:zC101N13.8;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kimberley A.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL772154; CAD61094.1; -
 DR InterPro: IPR009263; SERTA.
 DR Pfam: PF06031; SERTA; 1.
 SQ SEQUENCE 383 AA; 40402 MW; 76514201CDFC992F CRC64;

Query Match 12.9%; Score 133; DB 2; Length 383;
 Best Local Similarity 22.3%; Pred. No. 0.031;
 Matches 81; Conservative 35; Mismatches 71; Indels 176; Gaps 17;

QY 4 GLKRXHSLDEEEERWEW----SPAGLQ-----SY----QQALLRISLDKQVRSIGLGRAPS 51
 DB 27 GAKRK---LDEDEGLEKALAGAGAGLKVSYTLQRTIFNWSLMKLYNHRVTEPS 83
 QY 52 LRRHVLHNTLQQLQAALR-----PEPL-----FLGEEDFSLA-----L 71
 DB 84 LEKRVLLNNMLRRIQDELKQEGNLRPLFFPPPPDDPDVDSFREPQAPFVLSMWVAPPI 143
 QY 72 APAPALP-----TIGSILRELD-----TSMDGTTEP-----QNPVTPL 119
 DB 144 SQSPALSASSLTPSSSGLSNPALEACLTAPLLEEDNVSLCTSPSPPLAPPAPTSRLSPS 203
 QY 94 ---TIGSILRELD-----TSMDGTTEP-----QNPVTPL 119
 DB 204 VARDFSSALDEIEELCPSPPLTATSPAGTSPSPQLCPPSLNSGALDSKDCSKPCSPK 263
 QY 120 --GL-----QNEVPPQPD---PVFLE-----ALSSRYLGDSGLDD--FFLDIDTSVAVEK 161
 DB 264 LEGVLPLAERSAVPNTPETLPNLSLDMSTSPSASSSGFLTDALDDILFADIDTSMYDF 323
 QY 162 EP-----ARAPPEPPHNLFCAPGSWENWELDHIMEII 193
 DB 324 DPCTSSSGAAPSKLAPMVTADLLKTFSPYSGAAPAVSSN---QPFKMDLTDLHIMEVL 380
 QY 194 LGS 196
 DB 381 VGS 383

Search completed: November 15, 2004, 14:06:37
 Job time : 200 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2004, 14:06:47 ; Search time 142 Seconds
(without alignments)
488.370 Million cell updates/sec

Title: US-10-069-386-2

Perfect score: 1033

Sequence: 1 MEGGLRKHSJLEEEERWE.....APGSWEWELDHIMEILGS 196

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1568699 segs, 353819137 residues

Total number of hits satisfying chosen parameters: 1568699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pep.*
17: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	791	76.6	157	14	US-10-106-698-5516 Sequence 5516, Ap
2	645.5	62.5	142	14	US-10-029-386-33709 Sequence 33709, A
3	402	38.9	111	14	US-10-029-386-33708 Sequence 33708, A
4	216	20.9	236	11	US-09-978-360A-568 Sequence 568, App
5	213	20.6	236	9	US-09-731-872-412 Sequence 412, App
6	213	20.6	236	10	US-09-876-997-412 Sequence 412, App
7	213	20.6	222	15	US-10-296-115-1065 Sequence 1065, Ap
8	205	19.8	222	14	US-10-029-386-34001 Sequence 34001, A
9	136	13.2	237	13	US-10-076-069-2 Sequence 2, Appli
10	125	12.1	314	14	US-10-170-385-31 Sequence 31, Appl
11	125	12.1	314	16	US-10-755-889-627 Sequence 627, App
12	106	10.3	254	9	US-09-925-302-596 Sequence 596, App
13	106	10.3	254	10	US-09-925-302-596 Sequence 596, App

14	106	10.3	254	14	US-10-106-698-4456	Sequence 4456, Ap
15	105	10.2	241	13	US-10-076-069-4	Sequence 4, Appli
16	104.5	10.1	740	15	US-10-276-774-2083	Sequence 2083, Ap
17	98	9.5	562	14	US-10-104-047-3214	Sequence 3214, Ap
18	97	9.4	803	14	US-10-369-493-2536	Sequence 2536, Ap
19	97	9.4	1099	17	US-10-425-115-195088	Sequence 195088, A
20	96	9.3	297	15	US-10-425-114-64002	Sequence 64002, A
21	96	9.3	817	14	US-10-369-493-1813	Sequence 1813, Ap
22	95.5	9.2	1228	16	US-10-437-963-188300	Sequence 188300, A
23	95	9.2	278	17	US-10-739-930-8565	Sequence 8565, Ap
24	95	9.2	713	16	US-10-437-963-137248	Sequence 137248, A
25	94	9.1	433	8	US-08-259-451-3	Sequence 3, Appli
26	94	9.1	433	14	US-10-324-999A-3472	Sequence 3472, Ap
27	93.5	9.1	256	14	US-10-421-138A-312	Sequence 312, App
28	93.5	9.1	256	15	US-10-374-780A-1247	Sequence 1247, Ap
29	93.5	9.1	663	16	US-10-755-889-480	Sequence 480, App
30	93.5	9.1	1734	9	US-09-862-027-81	Sequence 81, Appl
31	93.5	9.1	1734	15	US-10-042-865-82	Sequence 82, Appl
32	93	9.0	157	15	US-10-076-701-54984	Sequence 54984, A
33	92.5	9.0	113	13	US-10-076-069-6	Sequence 6, Appli
34	92	8.9	1103	16	US-10-437-963-166325	Sequence 166325, A
35	91.5	8.9	714	15	US-10-425-114-62753	Sequence 62753, A
36	91	8.8	316	15	US-10-425-114-56782	Sequence 56782, A
37	91	8.8	690	17	US-10-425-115-221982	Sequence 221982, A
38	90.5	8.8	206	16	US-10-437-963-187376	Sequence 187376, A
39	90.5	8.8	356	17	US-10-425-115-257033	Sequence 257033, A
40	90.5	8.8	370	14	US-10-192-381-16	Sequence 16, Appl
41	90.5	8.8	824	14	US-10-226-844-1	Sequence 1, Appli
42	90.5	8.8	824	14	US-10-210-951-58	Sequence 58, Appl
43	90.5	8.8	824	14	US-10-211-884-58	Sequence 58, Appl
44	90.5	8.8	824	14	US-10-211-858-58	Sequence 58, Appl
45	90.5	8.8	947	9	US-09-871-889-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-10-106-698-5516
; Sequence 5516, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5516
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (55)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (132)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-5516

Query Match 76.6%; Score 791; DB 14; Length 157;

Best Local Similarity 98.0%; Pred. No. 1.3e-63;

Matches 150; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 44 SLGRAPSLRRHVLHNTLQQLQAALRLAPALPPEPLFGLGEDFSLSATIGSILRELD 103

```
DB 5 ALGPRAPSLRRHVLHNTLQQLQAALRLAPALPPPEFLGSEDFSLSAIGSILRELD 64
OY 104 TSMGTEPPQNVTPPLGLQNEVPQPPVFLEALSSRYLGDGLDDFFLDIDTSAYEKEP 163
DB 65 TSMGTEPPQNVTPPLGLQNEVPQPPVFLEALSSRYLGDGLDDFFLDIDTSAYEKEP 124
OY 164 ARAPPEPHNLFCAPGSWENNELDHIMEIILGS 196
DB 125 ARAPPEPHNLFCAPGSWENNELDHIMEIILGS 157

RESULT 2
US-10-029-386-33709
Sequence 33709, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
FILE REFERENCE: AROMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
SEQ ID NO 33709
LENGTH: 142
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC010271.5
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.64
US-10-029-386-33709

Query Match 62.5%; Score 645.5; DB 14; Length 142;
Best Local Similarity 96.8%; Pred. No. 1.7e-50;
Matches 122; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

OY 71 LAPALPPPEFLGSEDFSLSATIGSILRELTSMGTEPPQNVTPPLGLQNEVPQPD 130
DB 18 LPPLPC-PPEFLGSEDFSLSATIGSILRELTSMGTEPPQNVTPPLGLQNEVPQPD 76
OY 131 PVFLEALSSRYLGDGLDDFFLDIDTSAYEKEPARAPPEPHNLFCAPGSWENNELDHIM 190
DB 77 PVFLEALSSRYLGDGLDDFFLDIDTSAYEKEPARAPPEPHNLFCAPGSWENNELDHIM 136
OY 191 EILGS 196
DB 137 EILGS 142

RESULT 3
US-10-029-386-33708
Sequence 33708, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
FILE REFERENCE: AROMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
SEQ ID NO 33708
LENGTH: 111
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
```

```
; OTHER INFORMATION: MAP TO AC010271.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.64
US-10-029-386-33708

Query Match 38.9%; Score 402; DB 14; Length 111;
Best Local Similarity 89.9%; Pred. No. 1.4e-28;
Matches 80; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 MEGGLKRRKHSDEEEERWESWSPAGLQSYQQALLRISLDKQVRSGLGRAPSLRRHVLHN 60
DB 3 MVGGLKRRKHSDEEEERWESWSPAGLQSYQQALLRISLDKQVRSGLGRAPSLRRHVLHN 62
OY 61 TLQQLQAALRLAPALPPPEFLGSEDF 89
DB 63 TLQQLQAALRLAPALPPPEFLGSGGF 91

RESULT 4
US-09-978-360A-568
Sequence 568, Application US/09978360A
Publication No. US20040110939A1
GENERAL INFORMATION:
APPLICANT: Edwards, Jean-Baptiste Dumas Milne
APPLICANT: Duclert, Aymeric
APPLICANT: Bougueleret, Lydie
APPLICANT: Jobert, Severin
APPLICANT: Clusel, Catherine
TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
FILE REFERENCE: 56.US4.CIP
CURRENT APPLICATION NUMBER: US/09/978,360A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: US 60/066,677
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: US 60/069,957
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: US 60/074,121
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: US 60/081,563
PRIOR FILING DATE: 1998-04-13
PRIOR APPLICATION NUMBER: US 60/096,116
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: US 60/099,273
PRIOR FILING DATE: -09-04
PRIOR APPLICATION NUMBER: US 09/191,997
PRIOR FILING DATE: 1998-11-13
PRIOR APPLICATION NUMBER: US 09/215,435
PRIOR FILING DATE: 1998-12-17
PRIOR APPLICATION NUMBER: PCT/IB98/02122
PRIOR FILING DATE: 1998-12-17
PRIOR APPLICATION NUMBER: US 09/247,155
PRIOR FILING DATE: 1999-02-09
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 810
SOFTWARE: Patent.pm
SEQ ID NO 568
LENGTH: 236
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -31...-1
FEATURE:
NAME/KEY: misc.feature
LOCATION: (28, 30, 40, 67, 86, 117, 120)
OTHER INFORMATION: unknown
US-09-978-360A-568

Query Match 20.9%; Score 216; DB 11; Length 236;
Best Local Similarity 32.6%; Pred. No. 2.4e-11;
Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

OY 4 GLKRRKHSDEEEERWESWSPAGLQSYQQALLRISLDKQVRSGLGRAPSLRRHVLHN 51
```



```

; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 412
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-876-997-412

Query Match          20.6%; Score 213; DB 10; Length 236;
Best Local Similarity 32.6%; Pred. No. 4.6e-11;
Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

Qy  4 GLKRKHSDLLEE-----ERWESPAQLSQVQA-----LLRISLDKVSRLGPRAPS 51
      ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db  5 GLKRKREEEBEKEPLAVDSW-WLDPGHTVAQAAPPVAVASSFLDLSVLKHLHSLQQSEPD 63
      ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Qy  52 LRRHLVHNTLQOALRIAPALPPEPL-----FLGEEDFSLSATIGSILRE 101
      ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db  64 LRLHLVVNTRRIQAS--MAPAAALPVPSPPAAPSVADNLLASSDAALSASMASLLED 121
      ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Qy  102 LDTSMGDTEPPQNVTPLGLQNEVPP-----QPDVFLREAL--SSRYLGDSGLDDFFL 152
      ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db  122 L-SHIEGLSQAPQ-----LADEGPGRSIGGAAPSLGALDILGATGCLLDGGLGLEIFE 175
      ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Qy  153 DITSAVEKE---PARAPPPE-PHNLFCAPGSWE-----WNELDHIMEILGS 196
      ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db  176 DITSMYDNLWAPASEGLKPGPED---GPGKEAPELDEALDYLMDVLVGT 225
      ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

RESULT 7
US-10-296-115-1065
; Sequence 1065, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 794PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 1065
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-296-115-1065

Query Match          20.6%; Score 213; DB 15; Length 278;
Best Local Similarity 32.6%; Pred. No. 5.6e-11;
Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13

Qy  4 GLKRKHSDLLEE-----ERWESPAQLSQVQA-----LLRISLDKVSRLGPRAPS 51
      ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db  47 GLKRKREEEBEKEPLAVDSW-WLDPGHAAVAQAAPPVAVASSFLDLSVLKHLHSLQQSEPD 105
      ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Qy  52 LRRHLVHNTLQOALRIAPALPPEPL-----FLGEEDFSLSATIGSILRE 101
      ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db  106 LRLHLVVNTRRIQAS--MAPAAALPVPSPPAAPSVADNLLASSDAALSASMASLLED 163
      ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Qy  102 LDTSMGDTEPPQNVTPLGLQNEVPP-----QPDVFLREAL--SSRYLGDSGLDDFFL 152
      ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db  164 L-SHIEGLSQAPQ-----LADEGPGRSIGGAAPSLGALDILGATGCLLDGGLGLEIFE 217
      ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Qy  153 DITSAVEKE---PARAPPPE-PHNLFCAPGSWE-----WNELDHIMEILGS 196
      ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

```

DB 218 DDTSMYDNELWAPASEGLKXGPEP---GPGKEAPELDEALDYLMVDLVGT 267

RESULT 8

US-10-029-386-34001

Sequence 34001, Application US/10023386

Publication No. US20030194704A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

FILE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: EXPRESSION ANALYSIS TWO

FILE REFERENCE: AECOMICA-X-2

CURRENT APPLICATION NUMBER: US/10/029,386

CURRENT FILING DATE: 2001-12-20

NUMBER OF SEQ ID NOS: 34288

SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

SEQ ID NO 34001

LENGTH: 222

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AC010271.5

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.46

OTHER INFORMATION: SWISSPROT HIT: Q14140, EVALUATE 2.70e-01

US-10-029-386-34001

Query Match 19.8%; Score 205; DB 14; Length 222;
Best Local Similarity 32.6%; Pred. No. 2.2e-10;
Matches 75; Conservative 34; Mismatches 69; Indels 52; Gaps 13;

QY 4 GLKRXHSDLEEE---ERWESAPGLQSYQQA-----LRLSLDKVQSLGRAPAS 51
DB 5 GLKRXHSDLEEE---ERWESAPGLQSYQQA-----LRLSLDKVQSLGRAPAS 51
QY 52 LRRHLVHTLQQLQAALRALAPALPPEL-----FLGEEFSLSATIGSILRE 101
DB 64 LRLHLVHTLQQLQAALRALAPALPPEL-----FLGEEFSLSATIGSILRE 101
QY 102 LDTSMGTEPPQNVTEFLQNEVPP-----QPDVFLEAL--SSRYLGDSGLDDFFL 152
DB 122 L-SHIEGLSQAPQ-----LADEGPGKSTGGAPSLGALDLPATGCLLDGGLGLE 175
QY 153 DDTSAVEKE---PARAPPPE-PHNLFCAPGSWE-----WNELDHIMEII 193
DB 176 DDTSMYDNELWAPASEGLKXGPEP---GPGKEAPELDEALDYLMVDLV 222

RESULT 9

US-10-076-069-2

Sequence 2, Application US/10076069

Publication No. US20020177214A1

GENERAL INFORMATION:

APPLICANT: JURECIC, ROLAND

APPLICANT: NACHTMAN, RONALD

TITLE OF INVENTION: HEPP, A NOVEL GENE WITH A ROLE IN HEMATOPOIETIC AND NEURAL DEVELOPMENT

FILE REFERENCE: 39532-176599

CURRENT APPLICATION NUMBER: US/10/076,069

CURRENT FILING DATE: 2002-02-15

PRIOR APPLICATION NUMBER: US 60/268,923

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 11

SOFTWARE: PatentIn version 3.0

SEQ ID NO 2

LENGTH: 237

TYPE: PRT

ORGANISM: Mus musculus

US-10-076-069-2

Query Match 13.2%; Score 136; DB 13; Length 237;
Best Local Similarity 26.1%; Pred. No. 0.00043;
Matches 63; Conservative 32; Mismatches 86; Indels 60; Gaps 9;

QY 4 GLKRXHSDLEEEERWESPA--GLQSYQQAALRLSLDKVQSLGRAPASLRRHVLHNTL 62
DB 5 GLKRXHSDLEEEERWESPA--GLQSYQQAALRLSLDKVQSLGRAPASLRRHVLHNTL 62
QY 63 QQLQAALR-----LAP--APALPPELFLGEEFSLSATIGSILRLDTSMDGTEPPQ 113
DB 63 RQIQEWSQDGVHGMFQNDRAVERLVSTE---ILCRTVRGAEEHHPAPEDEAPLQ 119
QY 114 NVTPLGLQNEVPPQPDPP-----VFLGALSRYLGDSGLDDFF 151
DB 120 NSVSELPVGSAPGNGRNFQSSILWEMDSQENRGSFQKSLDQIFETLENK--NSSSVLELF 177
QY 152 LDTSAVEKEPAR-----APPEPHNLFCAPGSWEWNLHIMEII 193
DB 178 SDVDSYVDLDTVLTMGWSGTSKSLCNGLEGFAAATPPSPSTCKS---DLAELDHVVEIL 234
QY 194 L 194
DB 235 V 235

RESULT 10

US-10-170-385-31

Sequence 31, Application US/10170385

Publication No. US20030203372A1

GENERAL INFORMATION:

APPLICANT: Ward, Neil Raymond

APPLICANT: Mundy, Christopher Robert

APPLICANT: Kan, On

APPLICANT: Harris, Robert Alan

APPLICANT: White, Jonathan

APPLICANT: Binley, Katie Mary

APPLICANT: Rayner, William Nigel

APPLICANT: Naylor, Stuart

APPLICANT: Kingsman, Susan Mary

APPLICANT: Krige, David

TITLE OF INVENTION: ANALYSIS METHOD

FILE REFERENCE: 532682000100

CURRENT APPLICATION NUMBER: US/10/170,385

CURRENT FILING DATE: 2002-06-12

PRIOR APPLICATION NUMBER: PCT/GB02/01662

PRIOR FILING DATE: 2002-04-08

PRIOR APPLICATION NUMBER: PCT/GB01/05458

PRIOR FILING DATE: 2001-12-10

NUMBER OF SEQ ID NOS: 549

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 31

LENGTH: 314

TYPE: PRT

ORGANISM: Homo Sapiens

US-10-170-385-31

Query Match 12.1%; Score 125; DB 14; Length 314;

Best Local Similarity 22.7%; Pred. No. 0.006;

Matches 71; Conservative 34; Mismatches 58; Indels 120; Gaps 12;

QY 2 EGGKRXHSDLEEEERWESPA--GLQSYQQAALRLSLDKVQSLGRAPASLRRH 55
DB 4 KGG-KRKFDEHEDGLECKIVSPCDGPKSVSYTLQRTIFNLSMLKLYNHRPLTEPSLQKT 62
QY 56 VLIHNTLQQLQAALR-----YQALLRLSLDKVQSLGRAPASLRRH 70
DB 63 VLINMLRRIQELKQESLRPMFTPTSPQPTPTSDSYREAPPAPFSLGSPSHPCDGLS 122
QY 71 -----LAPA-----PALPPS---PLFLGEDFSLSATIGSILRELD 103
DB 123 TTPLEACLTFAILEDDEDDTCTSCAMQPTAPTCLSPALLPEKD-SFSSALDEIEELCP 181
QY 104 TSMDE-----GTEPPQNPVTEFLQNEVPPQ-----PDPVFLEAL-----SSRYLGDS 145
DB 182 TSTSTAATAATDSVKGTSEAGTQKLDGPGESRADDSKLMDSLPGNFEITTTSTGFTLTL 241

```
QY 146 GLDD-FFLDIDTSAVEKEP-----ARAPPEPHNLF-----CAPGSWEH 183
Db 242 TLDDILFADIDTSMYDFDCTSSGTASKAPVSAADDLLKTLAPYSSQPVTPSPQPKMDL 301
QY 184 NELDHIMEIILGS 196
Db 302 TELDHIMEVLVGS 314

RESULT 11
US-10-755-889-627
; Sequence 627, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 627
; TYPE: PRT
; LENGTH: 314
; ORGANISM: Homo sapiens
US-10-755-889-627

Query Match 12.1%; Score 125; DB 16; Length 314;
Best Local Similarity 22.7%; Pred. No. 0.006;
Matches 71; Conservative 34; Mismatches 88; Indels 120; Gaps 12;

QY 2 EGGLEKXKHSLEEEERWESWSPAGLQ-----YQALLRISLDKVSRLGPRAPSLRRH 55
Db 4 KGG-KRKFDEHEDGLEKIVSPCDGPKSVKSYTLQRTIFNISLMKLYNHRPLTEPSLQKT 62
QY 56 VLHNTLQQLAALR-----
Db 63 VLINMLRRIQEEELKQBSLRPMTSPSQTTSPDSYREAPPAPSHLASPSHPHCDLGS 122
QY 71 -----LAPA-----PALPPE---PLFLGDEDFSLSATIGSILRELD 103
Db 123 TTPLEACLTASLLEDDDDTFCFSQAMQPTAPTKLSPALLPEKD-SFSSALDEIEELCP 181
QY 104 TSMD-----GTPEPPQNPVTPGLQNEVPQO-----PDVPEAL-----SSRYLGDS 145
Db 182 TSTSTEAATAATDSVKGTSSEAGTQKLDGPQESRADDKSLMDSLPGNFETTTSTGFLTDL 241
QY 146 GLDD-FFLDIDTSAVEKEP-----ARAPPEPHNLF-----CAPGSWEH 183
Db 242 TLDDILFADIDTSMYDFDCTSSGTASKAPVSAADDLLKTLAPYSSQPVTPSPQPKMDL 301
QY 184 NELDHIMEIILGS 196
Db 302 TELDHIMEVLVGS 314

RESULT 12
US-09-925-302-596
; Sequence 596, Application US/09925302
; Patent No. US2002094491A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
```

```
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 596
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (105)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-596

Query Match 10.3%; Score 106; DB 9; Length 254;
Best Local Similarity 25.6%; Pred. No. 0.24;
Matches 65; Conservative 32; Mismatches 75; Indels 82; Gaps 13;

QY 4 GLKXKHSLEEEERWESWSPAGLQ---SY---QOALLRISLDKVSRLGPRAPSLRRHVL 57
Db 18 GLKXKCVGHEEDV---EGALAGLKTSSYSLSQRLDMSLVKLQCHMLVEPNLCRSVL 74
QY 58 IHNTLQQLAAL-----RLAP--APALPPEPLFLGDEDFSLSATIGSILRELDTSMG 108
Db 75 IANTVRQIEEMTQDGTWRTVAPQAEKAPXDRLVST-----ILCRAWGOEG 123
QY 109 TEP-----PQNPVT-----PLGLQNEVPQPD-----VFLEALS 138
Db 124 AHPAPCLGDGHTQGPVSLCPVTSQAQAPRHQSSAWMDGPRENKGSPHKSLDQIFETLE 183
QY 139 SRYLGDSGLDFFLDIDTSAVEKEP-----AR-----APPEPHNLCAPGS 180
Db 184 TK--NPSCEBELFSDVSPYYDLDTVLTMGMGAPGPGCEGLGLAPATPGFSSCKS-- 239
QY 181 WEWNELDHIMEIIL 194
Db 240 -DLGELDHVVEILV 252

RESULT 13
US-09-925-302-596
; Sequence 596, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 596
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (105)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-596

Query Match 10.3%; Score 106; DB 10; Length 254;
Best Local Similarity 25.6%; Pred. No. 0.24;
Matches 65; Conservative 32; Mismatches 75; Indels 82; Gaps 13;

QY 4 GLKXKHSLEEEERWESWSPAGLQ---SY---QOALLRISLDKVSRLGPRAPSLRRHVL 57
Db 18 GLKXKCVGHEEDV---EGALAGLKTSSYSLSQRLDMSLVKLQCHMLVEPNLCRSVL 74
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2004, 13:58:11 ; Search time 40 Seconds
(without alignments)
324.958 Million cell updates/sec

Title: US-10-069-386-2
Perfect score: 1033
Sequence: 1 MEGGLKXKHSLEEEERWE.....APGSNEWELDHIMBIILGS 196
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A.COMB.pcp.*
2: /cgn2_6/prodata/1/iaa/5B.COMB.pcp.*
3: /cgn2_6/prodata/1/iaa/6A.COMB.pcp.*
4: /cgn2_6/prodata/1/iaa/6B.COMB.pcp.*
5: /cgn2_6/prodata/1/iaa/6C.COMB.pcp.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	94	9.1	433	US-08-259-451-3	Sequence 3, Appli
2	93	9.0	433	US-07-672-483-4	Sequence 4, Appli
3	93	9.0	2618	US-09-413-814-28	Sequence 28, Appl
4	92.5	9.0	432	PCT-US95-04910-13	Sequence 13, Appl
5	90.5	8.8	370	US-09-377-285B-16	Sequence 16, Appl
6	90.5	8.8	947	US-08-887-518-2	Sequence 2, Appli
7	90.5	8.8	947	US-09-023-321-2	Sequence 2, Appli
8	90.5	8.8	947	US-09-032-475-2	Sequence 2, Appli
9	90.5	8.8	947	US-09-257-703-1	Sequence 1, Appli
10	90.5	8.8	947	US-09-871-889A-1	Sequence 1, Appli
11	89.5	8.7	215	US-08-778-717-9	Sequence 9, Appli
12	88.5	8.6	897	US-09-849-602-18	Sequence 18, Appl
13	88	8.5	1130	US-09-976-594-280	Sequence 280, App
14	87.5	8.5	709	US-10-118-328-4	Sequence 4, Appli
15	86.5	8.4	136	US-08-259-451-5	Sequence 5, Appli
16	84	8.1	337	US-09-543-681A-7444	Sequence 7444, Ap
17	84	8.1	514	US-09-252-991A-25281	Sequence 25281, A
18	84	8.1	628	US-09-345-473E-48	Sequence 48, Appl
19	84	8.1	2152	US-09-036-987A-3	Sequence 3, Appli
20	84	8.1	2152	US-09-370-700-3	Sequence 3, Appli
21	84	8.1	2152	US-09-603-207-3	Sequence 3, Appli
22	83	8.0	969	US-08-548-159-1	Sequence 1, Appli
23	83	8.0	985	US-08-548-159-3	Sequence 3, Appli
24	83	8.0	1012	US-08-811-481-16	Sequence 16, Appl
25	83	8.0	1012	US-08-876-527-16	Sequence 16, Appl
26	82.5	8.0	1130	US-09-538-092-834	Sequence 834, App
27	82.5	8.0	2142	US-09-538-092-1142	Sequence 1142, App

28	82	7.9	667	4	US-09-248-796A-18663	Sequence 18663, A
29	82	7.9	1006	4	US-09-023-905A-12	Sequence 12, Appl
30	82	7.9	2101	1	US-08-466-390-4	Sequence 4, Appli
31	82	7.9	2101	1	US-08-470-390-4	Sequence 4, Appli
32	82	7.9	2101	1	US-08-467-781-4	Sequence 4, Appli
33	82	7.9	2101	1	US-08-195-487-4	Sequence 4, Appli
34	82	7.9	2101	2	US-08-483-924-4	Sequence 4, Appli
35	82	7.9	2101	3	US-09-452-294-1	Sequence 1, Appli
36	82	7.9	2101	5	PCT-US93-06160-4	Sequence 4, Appli
37	81.5	7.9	1597	3	US-09-423-890-13	Sequence 13, Appl
38	81.5	7.9	1597	3	US-08-628-829-14	Sequence 14, Appl
39	81	7.8	357	4	US-09-252-991A-26979	Sequence 26979, A
40	81	7.8	535	3	US-08-813-574-2	Sequence 2, Appli
41	80.5	7.8	513	4	US-09-732-025-2	Sequence 2, Appli
42	80.5	7.8	583	4	US-09-252-991A-32735	Sequence 32735, A
43	80.5	7.8	804	4	US-09-270-767-46750	Sequence 46750, A
44	80.5	7.8	1162	2	US-08-728-323A-2	Sequence 2, Appli
45	80.5	7.8	1162	3	US-09-298-568-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-259-451-3
; Sequence 3, Application US/08259451
; Patent No. 6406841
; GENERAL INFORMATION:
; APPLICANT: Lee, Helen H.
; APPLICANT: Swanson, Priscilla A.
; APPLICANT: Idler, Kenneth B.
; APPLICANT: Rosenblatt, Joseph D.
; APPLICANT: Chen, Irvin S. Y.
; APPLICANT: Golde, David W.
; APPLICANT: Robertson, Eugene F.
; APPLICANT: Stephens, John E.
; APPLICANT: Chan, Emerson W.
; APPLICANT: Buytendorp, Mark H.
; APPLICANT: Johnson, Joan E.
; APPLICANT: Motley, Cheryl T.
; APPLICANT: Peterson, Bryan
; APPLICANT: Edwards, Michelle
; APPLICANT: Guidinger, Peggy
; APPLICANT: Tate, Cynthia
; TITLE OF INVENTION: HTLV-III/RA Compositions
; TITLE OF INVENTION: and Assays for Detecting HTLV Infection
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: One Abbott Park Road
; CITY: Abbott Park
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60064
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; MEDIUM TYPE: storage
; COMPUTER: IBM
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/259,451
; FILING DATE: 20-JUN-1994
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/086,415
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Daniel W. Collins
; REGISTRATION NUMBER: 31,912
; REFERENCE/DOCKET NUMBER: 5381.US.PI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (708) 937-6365

TELEFAX: (708) 938-2623

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 433 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

US-08-259-451-3

Query Match

Best Local Similarity 9.1%; Score 94; DB 4; Length 433;

Matches 40; Conservative 24; Mismatches 67; Indels 42; Gaps 7;

QY 47 PRAPSLRRHVLHNTLQLOALRLAPAPA-----LPPEPLFGBEDFSLSA 93

DB 13 PKAP---RGLSTHNLWFLQAAVRLQPSDFDFQQLRRFLKALKTPINLNPIDYSLLA 69

QY 94 TI-----GSILRELDTSMDGTEPPQNPVPLG---LQNEVPPQDPVFLFA-LSSRYLG 143

DB 70 SLIPKGYGPRVVEIINLVKNQVSPSAPAAVPTPICPTTTPPPPPSPGAHVPPPY-- 127

QY 144 DSGLLDFDLIDTSAVEKEPARAPPEPHNLFCAFGSWENWELDHIMEILGS 196

DB 128 -----VEPTTQCFFILHPPGAP-----SAHRPWQMKDLQAIKQEVSS 166

RESULT 2

US-07-672-483-4

Sequence 4, Application US/07672483

Patent No. 5359029

GENERAL INFORMATION:

APPLICANT: LACROIX, Martial

APPLICANT: ZREIN, Maan

TITLE OF INVENTION: PEPTIDES AND ANALOGUES AND MIXTURES

TITLE OF INVENTION: THEREOF FOR DETECTING ANTIBODIES TO HTLV-I AND HTLV-II

TITLE OF INVENTION: VIRUSES

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: FISH & NEAVE

STREET: 875 Third Avenue

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10022-6250

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/672,483

FILING DATE: 19910302

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/554,258

FILING DATE: 18-JUL-1990

ATTORNEY/AGENT INFORMATION:

NAME: HALEY Jr, James F

REGISTRATION NUMBER: 27,794

REFERENCE/DOCKET NUMBER: IAF8 CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 715-0600

TELEFAX: (212) 715-0674

TELEX: 14-8367

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 433 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-672-483-4

Query Match

Best Local Similarity 9.0%; Score 93; DB 1; Length 433;

Matches 40; Conservative 24; Mismatches 67; Indels 42; Gaps 7;

QY 47 PRAPSLRRHVLHNTLQLOALRLAPAPA-----LPPEPLFGBEDFSLSA 93

DB 13 PKAP---RGLSTHNLWFLQAAVRLQPSDFDFQQLRRFLKALKTPINLNPIDYSLLA 69

QY 94 TI-----GSILRELDTSMDGTEPPQNPVPLG---LQNEVPPQDPVFLFA-LSSRYLG 143

DB 70 SLIPKGYGPRVVEIINLVKNQVSPSAPAAVPTPICPTTTPPPPPSPGAHVPPPY-- 127

QY 144 DSGLLDFDLIDTSAVEKEPARAPPEPHNLFCAFGSWENWELDHIMEILGS 196

DB 128 -----VEPTTQCFFILHPPGAP-----SAHRPWQMKDLQAIKQEVSS 166

RESULT 2

US-07-672-483-4

Sequence 4, Application US/07672483

Patent No. 5359029

GENERAL INFORMATION:

APPLICANT: LACROIX, Martial

APPLICANT: ZREIN, Maan

Best Local Similarity 23.1%; Pred. No. 0.82;

Matches 40; Conservative 24; Mismatches 67; Indels 42; Gaps 7;

QY 47 PRAPSLRRHVLHNTLQLOALRLAPAPA-----LPPEPLFGBEDFSLSA 93

DB 13 PKAP---RGLSTHNLWFLQAAVRLQPSDFDFQQLRRFLKALKTPINLNPIDYSLLA 69

QY 94 TI-----GSILRELDTSMDGTEPPQNPVPLG---LQNEVPPQDPVFLFA-LSSRYLG 143

DB 70 SLIPKGYGPRVVEIINLVKNQVSPSAPAAVPTPICPTTTPPPPPSPGAHVPPPY-- 127

QY 144 DSGLLDFDLIDTSAVEKEPARAPPEPHNLFCAFGSWENWELDHIMEILGS 196

DB 128 -----VEPTTQCFFILHPPGAP-----SAHRPWQMKDLQAIKQEVSS 166

RESULT 3

US-09-413-814-28

Sequence 28, Application US/09413814

Patent No. 6225064

GENERAL INFORMATION:

APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH

APPLICANT: Bristol-Myers Squibb, Co.

APPLICANT: Beyer, Stefan

APPLICANT: Bloeker, Helmut

APPLICANT: Brandt, Petra

APPLICANT: Cino, Paul M

APPLICANT: Dougherty, Brian A

APPLICANT: Goldberg, Steven L

APPLICANT: Hoffe, Gerhard

APPLICANT: Mueller, Joachim

APPLICANT: Reichenbach, Hans

TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or

FILE REFERENCE: PCT/US 99/23535

CURRENT APPLICATION NUMBER: US/09/413,814

CURRENT FILING DATE: 1999-10-07

EARLIER FILING DATE: 1998-10-09

NUMBER OF SEQ ID NOS: 107

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 28

LENGTH: 2618

TYPE: PRI

ORGANISM: Sorangium cellulosum

US-09-413-814-28

Query Match

Best Local Similarity 9.0%; Score 93; DB 3; Length 2618;

Matches 55; Conservative 18; Mismatches 79; Indels 64; Gaps 11;

QY 31 QALLR---ISLD-----KVQSLGPRAP---SLRRHVLHNTLQLOQA 67

DB 974 QALLRRAIGLDDEFFQAGNSFGLIRLHAKLESAGKSPFITDLFQHTSIRSQAEMLSG 1033

QY 68 ALRLAP-APALPPPLFGBEDFSLSATIGSILRELDTSMDGTEPPQ---NPVTPFLGLQN 123

DB 1034 SSVAPLAGAVQPAAAAQVASSAAKSPGERGAAATSSGLTAQPPPHRPTIAVIGLAG 1093

QY 124 EVPEQPD-PVFLEAL-----SSRYLGDGSLD-----DFFLD 153

DB 1094 RPPAAPOLDLAFLELTGRCGIRPFPSQAELRDEGLDANRIACHNYVPKGFLLDRADHF-D 1152

QY 154 IDTSAVEKEPAR-APPEPHNLFCAFGSWENWELDH 188

DB 1153 ADFFGIPPRDAEITDQIRLLLECC-----WNALEH 1183

RESULT 4

PCT-US95-04910-13

Sequence 13, Application PC/TUS9504910

GENERAL INFORMATION:

APPLICANT: The Government of the United

APPLICANT: States of America as represented
APPLICANT: by the Secretary, Department of
APPLICANT: Health and Human Services
TITLE OF INVENTION: ISOLATION AND
TITLE OF INVENTION: CHARACTERIZATION OF A NOVEL PRIMATE T-CELL
TITLE OF INVENTION: LYMPHOTROPIC VIRUS AND THE USE OF THIS VIRUS
TITLE OF INVENTION: OR COMPONENTS THEREOF IN DIAGNOSTIC ASSAYS
TITLE OF INVENTION: AND VACCINES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04910
FILING DATE: 21-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US08/231,526
FILING DATE: 22-APR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAM S. FEILER
REGISTRATION NUMBER: 26,728
REFERENCE/DOCKET NUMBER: 2026-4125PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421722
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acids
STRANDEDNESS: unknown
TOPOLOGY: unknown
PCT-US95-04910-13
Query Match 9.0%; Score 92.5; DB 5; Length 432;
Best Local Similarity 21.5%; Pred. No. 0.92;
Matches 40; Conservative 19; Mismatches 48; Indels 79; Gaps 7;
QY 47 PRAPSLRRHVLHNTLQQLQAALRLAPAPA-----LPPEPLFLGEEDFSLSA 93
Db 13 PKAP---RGLSTHWNFLQASTYRLOQPSDFQQLRRFLKALKTPIWLPIDYSLJA 69
QY 94 TI-----GSLRELDTSMGTPEPPQNVPTPLGLQNEVPPQDPVFLEALSSRYLGD 144
Db 70 SLIPKGYPGRTSEIINVLIHQASPTPPAPSLP-----EPANPPPL----- 111
QY 145 SGLDDFELDITSAVEKEPARAPPEP-----PHNLFCAPGSWENNE 185
Db 112 -----QQPS-APPEPHTPPPIETPPATHCLPLILHPHGAPSAHRPQMOKD 154
QY 186 LDHIME 191
Db 155 LQAIKQ 160
RESULT 5
US-09-377-285B-16
Sequence 16, Application US/09377285B
Patent No. 6720175
GENERAL INFORMATION:
APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
APPLICANT: WORLEY, Paul
APPLICANT: TU, Jian

APPLICANT: XIAO, Bo
APPLICANT: LEAHY, Daniel
APPLICANT: BENKEN, Jutta
APPLICANT: LANAHAN, Anthony
TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING HOMER 1b PROTEIN (AS AMENDED)
FILE REFERENCE: JHU1580-4
CURRENT APPLICATION NUMBER: US/09/377,285B
CURRENT FILING DATE: 1999-08-18
PRIOR APPLICATION NUMBER: US 60/138,426
PRIOR FILING DATE: 1999-06-10
PRIOR APPLICATION NUMBER: US 60/138,493
PRIOR FILING DATE: 1999-06-10
PRIOR APPLICATION NUMBER: US 60/138,494
PRIOR FILING DATE: 1999-06-10
PRIOR APPLICATION NUMBER: US 60/097,334
PRIOR FILING DATE: 1998-08-18
NUMBER OF SEQ ID NOS: 72
SOFTWARE: Patent in version 3.0
SEQ ID NO 16
LENGTH: 370
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-377-285B-16
Query Match 8.8%; Score 90.5; DB 4; Length 370;
Best Local Similarity 23.5%; Pred. No. 1.2;
Matches 43; Conservative 29; Mismatches 62; Indels 49; Gaps 8;
QY 33 LLRLSLDKV-ORSLSGPRAPSLRRHVLHNTLQQLQAALRLAPAPALPP-----EPPLFLG 85
Db 102 MVNHLKVARREIGTLA-----IVVRLPPSQKVIPPESLPPLTPYCKRPL--- 147
QY 86 EEDFSLSATIGSILRELDTSMGT-----BPQNVPTPLGLQNEVPPQ-DPVPFLEAL-- 137
Db 148 --NFACLDDVGHGVKDLSTQLSRTGTLSRKSIKAPATPASATLGRPPRIPEPVQLPAVPD 205
QY 138 -----SSRYLGDSDGLDDFELDITSAVEKEPARAPPEP-----PHNLFCAPG 179
Db 206 GKLSAASSVSLSAGSAGSAGIPQSKGVAPATPPPPPIAPVTPPPPPPLPAEIFLLPP 265
QY 180 SWE 182
Db 266 PME 268
RESULT 6
US-08-887-518-2
Sequence 2, Application US/08887518
Patent No. 5843721
GENERAL INFORMATION:
APPLICANT: Rothe, Mike
APPLICANT: Wu, Lin
TITLE OF INVENTION: NTK Proteins, Nucleic Acids and Methods
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/887,518
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627

QY 3 GGLKRGHSDLEEEERWESPAGLOSQOALLRISLDKVORSIGLGRAPSLRRHVLIHNTL 62
Db 659 GGLK---SPWRGEYKEPRHPFNQANYHQ-----TLHAQPRELSRAPGPRPAEETTGRA 710
QY 63 QOLQALRLAPALAPPE-----PLFLGEEDFSLSATIGSILRELDTSMDGTEPPQNPTV 117
Db 711 PKLQ-----PPLPPEPEPNKSPPLTLSKEE-----SGMWEPLPLSSLEPAPARNPSS 758
QY 118 PLGLQNEVPPQ-----PDPVFLEALSSRY-----LGDSDLDDFFLDIDTSAVEKEPA 164
Db 759 P-ERKATVPEQELQLEIEFLNSLSQPFSLERQEQILSCLSDLSLSDDS---EKNPS 814
QY 165 RAPPEPHNLFCAPGSW-----EWN 184
Db 815 KASQSSRDTLSSGVHSSQAEARSSWN 843
RESULT 9
US-09-257-703-1
; Sequence 1, Application US/09257703
; Patent No. 6265538
; GENERAL INFORMATION:
; APPLICANT: Greene, Warner C.
; APPLICANT: Lin, Xin
; APPLICANT: Gelezuinas, Romas
; TITLE OF INVENTION: A NOVEL INHIBITOR OF THE INFLAMMATORY RESPONSE INDUCED
; FILE OF INVENTION: BY TNF-ALPHA AND IL-1
; FILE REFERENCE: 30448.61USU1
; CURRENT APPLICATION NUMBER: US/09/257,703
; CURRENT FILING DATE: 1999-02-25
; EARLIER APPLICATION NUMBER: 60/076,299
; EARLIER FILING DATE: 1998-02-27
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 947
; TYPE: PRT
; ORGANISM: NF-KAPPA B INDUCING KINASE (NIK)
US-09-257-703-1
Query Match 8.8%; Score 90.5; DB 3; Length 947;
Best Local Similarity 25.8%; Pred. No. 4;
Matches 54; Conservative 17; Mismatches 87; Indels 51; Gaps 10;
QY 3 GGLKRGHSDLEEEERWESPAGLOSQOALLRISLDKVORSIGLGRAPSLRRHVLIHNTL 62
Db 659 GGLK---SPWRGEYKEPRHPFNQANYHQ-----TLHAQPRELSRAPGPRPAEETTGRA 710
QY 63 QOLQALRLAPALAPPE-----PLFLGEEDFSLSATIGSILRELDTSMDGTEPPQNPTV 117
Db 711 PKLQ-----PPLPPEPEPNKSPPLTLSKEE-----SGMWEPLPLSSLEPAPARNPSS 758
QY 118 PLGLQNEVPPQ-----PDPVFLEALSSRY-----LGDSDLDDFFLDIDTSAVEKEPA 164
Db 759 P-ERKATVPEQELQLEIEFLNSLSQPFSLERQEQILSCLSDLSLSDDS---EKNPS 814
QY 165 RAPPEPHNLFCAPGSW-----EWN 184
Db 815 KASQSSRDTLSSGVHSSQAEARSSWN 843
RESULT 10
US-09-871-889A-1
; Sequence 1, Application US/09871889A
; Patent No. 6645728
; GENERAL INFORMATION:
; APPLICANT: Greene, Warner C.
; APPLICANT: Lin, Xin
; APPLICANT: Gelezuinas, Romas
; TITLE OF INVENTION: A NOVEL INHIBITOR OF THE INFLAMMATORY RESPONSE INDUCED BY TNF-ALF
; FILE REFERENCE: 30448.61USU1

; CURRENT APPLICATION NUMBER: US/09/871,889A
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 09/257,703
; PRIOR FILING DATE: 1999-02-25
; PRIOR APPLICATION NUMBER: 60/076,299
; PRIOR FILING DATE: 1998-02-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 947
; TYPE: PRT
; ORGANISM: NF-KAPPA B INDUCING KINASE (NIK)
US-09-871-889A-1
Query Match 8.8%; Score 90.5; DB 4; Length 947;
Best Local Similarity 25.8%; Pred. No. 4;
Matches 54; Conservative 17; Mismatches 87; Indels 51; Gaps 10;
QY 3 GGLKRGHSDLEEEERWESPAGLOSQOALLRISLDKVORSIGLGRAPSLRRHVLIHNTL 62
Db 659 GGLK---SPWRGEYKEPRHPFNQANYHQ-----TLHAQPRELSRAPGPRPAEETTGRA 710
QY 63 QOLQALRLAPALAPPE-----PLFLGEEDFSLSATIGSILRELDTSMDGTEPPQNPTV 117
Db 711 PKLQ-----PPLPPEPEPNKSPPLTLSKEE-----SGMWEPLPLSSLEPAPARNPSS 758
QY 118 PLGLQNEVPPQ-----PDPVFLEALSSRY-----LGDSDLDDFFLDIDTSAVEKEPA 164
Db 759 P-ERKATVPEQELQLEIEFLNSLSQPFSLERQEQILSCLSDLSLSDDS---EKNPS 814
QY 165 RAPPEPHNLFCAPGSW-----EWN 184
Db 815 KASQSSRDTLSSGVHSSQAEARSSWN 843
RESULT 11
US-08-778-717-9
; Sequence 9, Application US/08778717
; Patent No. 6602689
; GENERAL INFORMATION:
; APPLICANT: UENO, EIICHI
; APPLICANT: NOBUYUKI, FUJII
; APPLICANT: OKADA, MASAHISA
; TITLE OF INVENTION: FUSED DNA SEQUENCE, FUSED PROTEIN
; TITLE OF INVENTION: EXPRESSED FROM SAID FUSED DNA SEQUENCE AND METHOD FOR
; TITLE OF INVENTION: EXPRESSING SAID DNA SEQUENCE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/778,717
; FILING DATE: 12-DEC-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 352225/1995
; FILING DATE: 28-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2084-031-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000

TELEFAX: 703-413-2220

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 215 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: RECOMBINANT

PUBLICATION INFORMATION:

AUTHORS: NORIYUKI FUJII ET AL.

TITLE: FUSED DNA SEQUENCE, FUSED PROTEIN EXPRESSED

TITLE: FROM SAID FUSED DNA SEQUENCE AND METHOD OF

TITLE: EXPRESSING SAID FUSED PROTEIN

RELEVANT RESIDUES IN SEQ ID NO: 9: FROM 1 TO 215

US-08-778-717-9

Query Match 8.7%; Score 89.5; DB 4; Length 215;
Best Local Similarity 23.7%; Pred. No. 0.72; Mismatches 67; Indels 41; Gaps 6;
Matches 41; Conservative 24;

Qy 2 EGGKRRKHSLEEE-----ERWSPAGLQSYQOALLRISLDKQVRSGLG----- 46
Dy 31 EGRAQPKVEIEDEELYNCAKEAMEACFVSAITIEBAGGSSILVPRGSEFMGQIHGLSPPT 90
Qy 47 -PRAPSLRRHLVHNTLQOQALRLAPAPA-----LPEPLFLGEEEDFSL 92
Dy 91 IPKAP---RGLSTHMLNLFQOAYRQPRSDDFQOLRRFLKLAUKTFIWLNPIDISLT 147
Qy 93 ATI-----GSILRELDTSMDGTEPPQNPVTPLG---LQNEVPPQDPVPVLEA 136
Dy 148 ASLIPKYGPRGVVEIINILVKNQVSPSAPAPVPTPICPTTTPPPPPSPSPA 200

RESULT 12

US-09-849-602-18

Sequence 18, Application US/09849602

Patent No. 6794501

GENERAL INFORMATION:

APPLICANT: Scanlan, Matthew J.

APPLICANT: Old, Lloyd J.

APPLICANT: Stockert, Elisabeth

APPLICANT: Chen, Yao-Tseng

TITLE OF INVENTION: Colon Cancer Antigen Panel

FILE REFERENCE: L0461/7105(JRV)

CURRENT APPLICATION NUMBER: US/09/849,602

CURRENT FILING DATE: 2001-05-04

NUMBER OF SEQ ID NOS: 30

SOFTWARE: Patent version 3.0

SEQ ID NO 18

LENGTH: 897

TYPE: PRT

ORGANISM: Homo sapiens

US-09-849-602-18

Query Match 8.8%; Score 88.5; DB 4; Length 897;
Best Local Similarity 24.4%; Pred. No. 5.7; Mismatches 50; Conservative 23; Indels 75; Gaps 9;
Matches 50;

Qy 13 EEEERWSPAGLQSYQO-----ALLRISLDKQVRSGLGRAP-----SL 52
Dy 208 QQEAER-----QALQSLRQGGTITGKEMSTSIIPGCLLGVLE-----GDSPHGHASL 256
Qy 53 RRHLVHNTLQOQALRLAPALPPELFLGEEEDFSLSATIGSLRELDTSMDGTPPT 112
Dy 257 LQHVLL---LEQAQQSTLITVPLHQQSQSLVITGERVATSMRTVKGKLPRLRLSR--- 312
Qy 113 QNPVTPGLQNEVPPQDPVPVLEALSSR---YLGDSGLDDFFLDITDSAVEKEPARAPPEP 170
Dy 313 -LPQSFQALQQLVQOQOCHQOFLQKQOQOLQK-----ILTKTGLFROPTTHPEET 364
Qy 171 PHNLFAPGSGWENELDHIMEIILG 195

Db 365 E-----BELTEQQEVLLG 377
||:|:|

RESULT 13

US-09-976-594-280

Sequence 280, Application US/09976594

Patent No. 6673549

GENERAL INFORMATION:

APPLICANT: Buchbinder, Jenny

APPLICANT: Furness, Michael

TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS

FILE REFERENCE: PA-0041 US

CURRENT APPLICATION NUMBER: US/09/976,594

CURRENT FILING DATE: 2001-10-12

PRIOR APPLICATION NUMBER: 60/240,409

PRIOR FILING DATE: 2000-10-12

NUMBER OF SEQ ID NOS: 1143

SOFTWARE: PERL Program

SEQ ID NO 280

LENGTH: 1130

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. 6673549 2258794CD1

US-09-976-594-280

Query Match 8.5%; Score 88; DB 4; Length 1130;
Best Local Similarity 21.8%; Pred. No. 8.6; Mismatches 52; Conservative 23; Indels 88; Gaps 11;
Matches 52;

Qy 5 LKRKHSLEEEERWSPAGL-----QSYQOALLRISLD----- 39
Dy 539 LKR-----LKAKEEWREARQQGPNKIWEQYKAYLK-SLDHQAQVNFQNDTKALRSKSL 593
Qy 40 -----KVQSLGPRAPSLRRH-VLIHNTLQOQALRLAP-----APALPPE--- 80
Dy 594 NEIESVYDEHQEQHSEGRSAPSPSEPHLIFVYEDRQILEDAAALISYVVKRQPAIQKEDQG 653
Qy 81 -----PLFLGEEEDFSLSATIGSLRELDTSMDGTEPPQNPVTPGLQNEVPPQDP 130
Dy 654 TIHOLLHQFVPSLFFSQOQLDGLASESADRDSPQOQTTPDSEKPKAPGPHSSPEEK 713
Qy 131 PVFLEALSSRYLGDSDGLDDFFLDITDSAVEKEPARAPPEPPH-----NLFCAPGSW 181
Dy 714 GAF-----GDA-----PATEQPP--LPPAPAPKPLDDVYSLFTANNW 749

RESULT 14

US-10-118-328-4

Sequence 4, Application US/10118328

Patent No. 6773904

GENERAL INFORMATION:

APPLICANT: YAN, Chunhua et al.

TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE

TITLE OF INVENTION: PROTEINS, AND USES THEREOF

FILE REFERENCE: CLC01220

CURRENT APPLICATION NUMBER: US/10/118,328

CURRENT FILING DATE: 2002-04-09

PRIOR APPLICATION NUMBER: 60/282,460

PRIOR FILING DATE: 2001-04-10

NUMBER OF SEQ ID NOS: 5

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 4

LENGTH: 709

TYPE: PRT

ORGANISM: Mus musculus

US-10-118-328-4

Query Match 8.5%; Score 87.5; DB 4; Length 709;
Best Local Similarity 26.6%; Pred. No. 5.3;

Matches 46; Conservative 18; Mismatches 62; Indels 47; Gaps 9;
QY 5 LKKKHSDBEEERWESWAGLQSQOALIRISLD-----KVQSLG----- 46
Db 440 LKGNLNFKEKKEWE-ILARIQOQRCQVSLSPRPILAALRAQRLSEEQSVYVSR 498
QY 47 -----PRAPSLRRHVLIHNTLO-QLOAALRLAP--APALPEPLFLGEEDFSLSAT 94
Db 499 VIEPPAASCPSPIRRIRISLTKRLSAKLSREKNSSPGSGPGDPSSP-----TSSVS 550
QY 95 IG---SILRELDTSMDGTBP---PONPVTPGLQNEVPQDPVFLALSSR 140
Db 551 PGSPSPSPRNRRPPPGSPASPSPGSPSTKLSLTD-PPGWPVTLTPSSSR 602

RESULT 15

US-08-259-451-5
; Sequence 5, Application US/08259451
; Patent No. 6408841
; GENERAL INFORMATION:
; APPLICANT: Lee, Helen H.
; APPLICANT: Swanson, Priscilla A.
; APPLICANT: Idler, Kenneth B.
; APPLICANT: Rosenblatt, Joseph D.
; APPLICANT: Chen, Irvin S. Y.
; APPLICANT: Golde, David W.
; APPLICANT: Robertson, Eugene F.
; APPLICANT: Stephens, John E.
; APPLICANT: Chan, Emerson W.
; APPLICANT: Buytendorp, Mark H.
; APPLICANT: Johnson, Joan E.
; APPLICANT: Motley, Cheryl T.
; APPLICANT: Peterson, Bryan
; APPLICANT: Edwards, Michelle
; APPLICANT: Guidinger, Peggy
; APPLICANT: Tate, Cynthia
; TITLE OF INVENTION: HTLV-IIINRA Compositions
; TITLE OF INVENTION: and Assays for Detecting HTLV Infection
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: One Abbott Park Road
; CITY: Abbott Park
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60064

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; MEDIUM TYPE: storage
; COMPUTER: IBM
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/259,451
; FILING DATE: 20-JUN-1994
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/086,415
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; .NAME: Daniel W. Collins
; REGISTRATION NUMBER: 31,912
; REFERENCE/DOCKET NUMBER: 5381.US.P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (708) 937-6365
; TELEFAX: (708) 938-2623
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 136 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown

US-08-259-451-5

Query Match 8.4%; Score 86.5; DB 4; Length 136;
Best Local Similarity 26.8%; Pred. No. 0.77;
Matches 30; Conservative 15; Mismatches 42; Indels 25; Gaps 4;
QY 47 PRAPSLRRHVLIHNTLOQLOAALRLAPAP-----LPPPEPLFLGEEDFSLSA 93
Db 13 PKAP---RGUSTHWNLFQAAYRLQPGSDFFDQQRRLFKALKATPIWLNPIYISLLA 69
QY 94 TI-----GSILRELDTSMDGTBPQNPVTPLG---LQNEVPPQDPVFLFA 136
Db 70 SLIPKGYPRVVEIINILVKNQVSFSAAPVPTPICTTTTPPPPPPPSPEA 121

Search completed: November 15, 2004, 14:08:14
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2004, 13:57:50 ; Search time 38 Seconds
(without alignments)
496.276 Million cell updates/sec

Title: US-10-069-386-2

Perfect score: 1033

Sequence: 1 MEGGLKRXHSDLEEEERWE.....AFGSWEWNELDHIMEILLGS 196

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101.5	9.6	728	2 H59435	phosphoinositide-3
2	97	9.4	803	2 T40514	Chaperonin hsp78p
3	97	9.4	1201	2 G86441	unknown protein [i
4	96	9.3	817	2 S51342	verprolin - yeast
5	95.5	9.2	446	2 A42029	transcription fact
6	93.5	9.1	1734	2 A54602	microtubule-associ
7	93	9.0	433	1 POLH2	gag polyprotein -
8	91.5	8.9	1621	2 A82255	hypothetical prote
9	90.5	8.8	804	2 AG0565	probable membrane
10	88	8.5	651	2 T42644	hypothetical prote
11	88	8.5	716	2 T26998	hypothetical prote
12	88	8.5	1676	2 A6508	anucleate primary
13	86.5	8.4	536	2 A34596	transcription fact
14	86.5	8.4	1494	2 T14355	protein-tyrosine p
15	86	8.3	475	2 B37761	ntira protein - Thi
16	85.5	8.3	485	2 T37550	hypothetical coile
17	85.5	8.3	825	2 E75508	conserved hypothet
18	85.5	8.3	1544	2 E59431	phosphoinositide-b
19	85	8.2	715	2 D85087	hypothetical prote
20	85	8.2	1217	2 T00270	hypothetical prote
21	84.5	8.2	366	1 SAVLWD	large surface anti
22	84.5	8.2	477	2 T46304	hypothetical prote
23	84.5	8.2	723	2 B38749	3-phosphatidylinos
24	84.5	8.2	908	2 T16057	hypothetical prote
25	84.5	8.2	982	2 T43676	hunchback-related
26	84.5	8.2	1262	2 T13353	protein stn-B - fr
27	84	8.1	289	2 D70452	leucine-tRNA ligas
28	84	8.1	954	2 I61714	co-repressor prote
29	84	8.1	1015	2 JC5062	phogrin precursor

30	84	8.1	1026	2 T28968	hypothetical prote
31	84	8.1	1234	2 T00363	hypothetical prote
32	84	8.1	1819	2 T32008	hypothetical prote
33	83.5	8.1	331	2 S09800	hypothetical prote
34	83.5	8.1	601	2 AG0086	probable AMP-bindi
35	83.5	8.1	804	2 A85549	probable oxidoredu
36	83.5	8.1	804	2 G90698	probable oxidoredu
37	83	8.0	346	2 E72672	hypothetical prote
38	83	8.0	737	2 S28030	DNA-binding protei
39	83	8.0	1353	2 T00249	ich1 protein - ink
40	82.5	8.0	555	2 T30349	structural protein
41	82.5	8.0	1130	1 TVRUA	protein-tyrosine k
42	82.5	8.0	1870	2 S37671	MHC class III hist
43	82.5	8.0	1872	2 S36152	MHC class III hist
44	82.5	8.0	2142	2 B35098	MHC class III hist
45	82	7.9	450	2 B47265	tailless (tll) pro

ALIGNMENTS

RESULT 1

H59435

phosphoinositide-3-kinase regulatory beta chain [imported] - human

C:Species: Homo sapiens (man)

C>Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 09-Jul-2004

C:Accession: H59435; A59436

R:Volinia S; Patracchini P; Otsu M; Hiles I; Gout I; Calzolari E; Bernardi F; Rooke L; V

Oncogene 7, 789-793, 1992

A:Title: Chromosomal localization of human p85 alpha, a subunit of phosphatidylinositol

A:Reference number: H59435

A:Accession: H59435

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-728 <VOL>

A:Cross-references: UNIPROT:O00459; GB:NP_005018; PID:g4826908; PIDN:NP_005018.1

R:Jaassen, J.W.; Schluthoff, L.; Bartram, C.R.; Schulz, A.S.

Oncogene 16, 1767-1772, 1998

A:Title: An oncogenic fusion product of the phosphatidylinositol 3-kinase p85beta subun

A:Reference number: A59436; MUID:98241181; PMID:9582025

A:Accession: A59436

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-728 <JAN>

A:Cross-references: GB:NP_005018; PID:g4826908; PIDN:NP_005018.1

Query Match 9.8%; Score 101.5; DB 2; Length 728;

Best Local Similarity 23.9%; Pred. No. 1;

Matches 47; Conservative 25; Mismatches 70; Indels 55; Gaps 9;

QY 20 EWSPA---GLQSYQAL-----LRSLDKVQSLGPRAPSLRRHVLH 59

Db 164 QMTHAALADGKSFLLALPAFLVTPESAARRALREAGVGALEPTPLHRAITLR 223

QY 60 NTLOQLAALRLAPALPEPLFLGEDFSLSATIGSILRELTSMGTPPPONVPTPL 119

Db 224 FLLOHLGRVARRAPA-----LGPVRAIGATFGPLL-----LRAPPPSPSP-PFG 267

QY 120 GLQNEVPFQPD--PVFLEALSSRYLGSDGLDDFLDITSAVEKEPARAPPEP----- 170

Db 268 GAPDGESEFPDPFALLVEKLQEHLEQ-----EVAPPALPPAPKAPFVLIANGG 320

QY 171 -PHNLFCAFGSWEWNEEL 186

Db 321 SPFSL--QDAEWYWGDI 335

RESULT 2

T40514

Chaperonin hsp78p - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: T40514

R.Gwilliam, R.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.
submitted to the EMBL Data Library, September 1998
A:Reference number: Z21933
A:Accession: T40514
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-803 <GMI>
A:Cross-references: UNIPROT:O74402; EMBL:AL031534; PIDN:CAA20737.1; GSPDB:GN000057; SPDB:
A:Experimental source: strain 972h-; cosmid c4f6
C:Genetics:
A:Gene: SPDB:SPBC4F6.17C
A:Map position: 2
C:Superfamily: endopeptidase Clp ATP-binding chain
C:Keywords: ATP; molecular chaperone; nucleotide binding; P-loop
F:139-146/Region: nucleotide-binding motif A (P-loop)
F:207-212/Region: nucleotide-binding motif B
F:539-546/Region: nucleotide-binding motif A (P-loop)
F:607-612/Region: nucleotide-binding motif B
F:145/Binding site: ATP (Lys) #status predicted
F:545/Binding site: ATP (Lys) #status predicted
Query Match 9.4%; Score 97; DB 2; Length 803;
Best Local Similarity 24.2%; Pred. No. 2.8;
Matches 53; Conservative 31; Mismatches 73; Indels 62; Gaps 11;
OY 5 LKRHSLSLEEEER-----NEWSPAGLQSYQQA-----LLRSLDKVQSRSLGPRA----- 49
DB 377 LESKLTULKEQDKLSAWEERKLLDSIKKATELEQARIELERTQRE-GNVARASELQ 435
OY 50 ----PSLRHV-----LIHTLQQLQALRLAPALPPEPLFLGEEB--RS 90
DB 436 YAIIPELERSVPKEKTELEKXPSMVHDSVTSDDIAVWSRATGIPITTLMRGERDKLLN 495
OY 91 LSATIGSLRELDTSMGTGPQPQNPVTPGLQNEVPPQPPVLEALSSRYLGDSG---- 146
DB 496 MEQTIGKKIKQDALKAIADAVR-LSRAGLQNTNP-----LASFLGLGTGVGKT 546
OY 147 ----LDDFFLDIDTSAV-----EKEP-ARAPPEPP 171
DB 547 ALTKALAEFLDTDKAMIRFDMSEFOEKHTIARLIGSPP 585
RESULT 3
G86441
Unknown protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: G86441
B:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
Angen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
C:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Pizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
Ker, M.; Wu, B.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A85141; MUID:21016719; PMID:11130712
A:Accession: G86441
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1201 <STO>
A:Cross-references: UNIPROT:Q9C6S1; GB:AE005172; MID:g11136725; PIDN:AAG31306.1; GSPDB:G
C:Genetics:
A:Map position: 1
Query Match 9.4%; Score 97; DB 2; Length 1201;
Best Local Similarity 24.6%; Pred. No. 4.7;
Matches 41; Conservative 11; Mismatches 53; Indels 62; Gaps 6;
OY 13 EEEERWESPAGLQSYQQAALLRISDKVQSRSLGPAPSLRRHVLHNTLQQLQALRLA 72

Db 439 SSSEPPHESH---HHHHEIFAKOSVDN-PLNLSDDPPSSGDHVTL-----LP 482
OY 73 PAPALPPEPLFLGEEBDFLSLATSIGSILRELTSMGTGPQPQNPVTPGLG-----ONE 124
DB 483 PEPPEPPPEPLTSTTSFSPS-----QPPEPPPEPLFMSTTSFSPSQPP 526
OY 125 VPQPDVPVLEALSRYLGDSGLDDFFLDIDTSAREKPEAPAPPEPP 171
DB 527 PEPPEPPPLF-----TSTTSFSPSQPPPEPP 551
RESULT 4
S51342
varprolin - yeast (Saccharomyces cerevisiae)
N:Alternate names: prolin-rich protein VRP1; protein LB300.13; protein YLR337C
C:Species: Saccharomyces cerevisiae
C:Date: 23-Feb-1995 #sequence_revision 11-Aug-1995 #text_change 09-Jul-2004
C:Accession: S51342; S39626; S57435
R:Du, Z.
submitted to the EMBL Data Library, December 1994
A:Description: The sequence of S. cerevisiae cosmid 8300.
A:Reference number: S51339
A:Accession: S51342
A:Molecule type: DNA
A:Residues: 1-817 <DUZ>
A:Cross-references: UNIPROT:P37370; EMBL:U19028; NID:G609380; PID:G609392; MIPS:YLR397G
R:Donnelly, S.F.H.; Pocklington, M.J.; Pallotta, D.; Orr, E.
Mol. Microbiol. 10, 585-596, 1993
A:Title: A proline-rich protein, verprolin, involved in cytoskeletal organization and ce
A:Reference number: S39626; MUID:95058201; PMID:7968536
A:Accession: S39626
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-307 'R', 309-349, 'R', 351-688, 'E', 690-709, 'HLRWIPVELIAPVKTLNNGYFLQVDRGNNISIG
A:Cross-references: EMBL:Z26645; NID:G414785; PIDN:CAA81388.1; PID:G439289
R:Munn, A.L.; Stevenson, S.D.; Geli, M.I.; Riezman, H.
submitted to the EMBL Data Library, June 1995
A:Description: ends, end6, and end7: mutations that cause actin delocalization and block
A:Reference number: S57435
A:Accession: S57435
A:Molecule type: DNA
A:Residues: 1-162, 'F', 164-817 <MUN>
A:Cross-references: EMBL:X87806; NID:G871534; PID:G871535
A:Experimental source: strain W303
C:Genetics:
A:Gene: SGD:VRP1; MDPS; ENDS
A:Cross-references: SGD:S0004329; MIPS:YLR337C
A:Map position: 12R
F:5-14/Region: proline-rich
F:77-85/Region: proline-rich
F:114-180/Region: proline-rich
F:216-245/Region: proline-rich
F:305-336/Region: proline-rich
F:349-357/Region: proline-rich
F:372-382/Region: proline-rich
F:396-406/Region: proline-rich
F:421-445/Region: proline-rich
F:518-528/Region: proline-rich
F:567-577/Region: proline-rich
F:608-621/Region: proline-rich
F:649-661/Region: proline-rich
F:678-685/Region: proline-rich
F:704-710/Region: proline-rich
Query Match 9.3%; Score 96; DB 2; Length 817;
Best Local Similarity 25.5%; Pred. No. 3.5;
Matches 38; Conservative 17; Mismatches 44; Indels 50; Gaps 7;
OY 39 DKVORSIG---PRAPSLRRHVLHNTLQQLQALRLAPALPPEPLFLGEEBDFLSLATS 95
DB 594 DKSTVLQKKAKAPALPGHV-----PPFPVP-----VLSDSKNNVPA 633
OY 96 GSILRELTSMGTGPQPQNPV-----TPGLQNEVP---PQDPVLEALSRYL 142

Db 634 ASLLHDVLPSSNLEKPPSPVAAPPLPTFSAESLQQSVSTSIPEPPV-APTLSVR-- 690
Qy 143 GDSGLDDFFLDIDTSAVEKEPARAPPEPP 171
Db 691 -----TETESIKNFTKSPPPPP 708

RESULT 5
A42029
transcription factor E3 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A42029
R:Roman, C.; Matera, A.G.; Cooper, C.; Artandi, S.; Blain, S.; Ward, D.C.; Calame, K.
Mol. Cell. Biol. 12, 817-827, 1992
A>Title: mTFE3, an X-linked transcriptional activator containing basic helix-loop-helix
A:Reference number: A42029; MUID:92123207; PMID:1732746
A:Accession: A42029
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-446 <ROM>
A:Cross-references: UNIPROT:Q64092; GB:S76673; NID:G243439; PIDN:AB21130.1; PID:G243440
A>Note: sequence extracted from NCBI backbone (NCBI:76673, NCBI:76674)

Query Match 9.2%; Score 95.5; DB 2; Length 446;
Best Local Similarity 27.3%; Pred. No. 1.7;
Matches 48; Conservative 24; Mismatches 61; Indels 43; Gaps 8;

Qy 12 LEEBERWSPAGLQSQQALLRISLDKVSRLGRAPSLRRHVLIHTLQQLAALRL 71
Db 272 LQEQQR-----SKDLSRQR-----SLEQANRSLQRIQEL-----ELQAIHG 311
Qy 72 APAPALPPEPLFGEEDFSLSATISILRELDTSMDGTEPPQNPVTPGLQNEVPPQPPD 131
Db 312 LP---VPPNGLSLTSSVSDSLKP--EQLDIEESGRPSTTHVSGGPAQNAPPQPPA 366

Qy 132 VFLEAL-----SRYGDSGLDDFFLDIDTSAVEKE-----PARAPPEP 170
Db 367 PPSDALLDLHFPBDHLGDLG-DPFLHGLDILMEEGVMVGGSLGGALSPIRAASDP 421

RESULT 6
A54602
microtubule-associated serine/threonine protein kinase MAST205 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: A54602
R:Walden, P.D.; Cowan, N.J.
Mol. Cell. Biol. 13, 7625-7635, 1993
A>Title: A novel 205-kilodalton testis-specific serine/threonine protein kinase associat
A:Reference number: A54602; MUID:94067123; PMID:8246979
A:Accession: A54602
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1734 <WAL>
A:Cross-references: UNIPROT:Q40592; GB:U02313; NID:G406057; PIDN:AAC04312.1; PID:G406058
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C:Keywords: ATP

F:451-726/Domain: protein kinase homology <XIN>
F:459-467/Region: protein kinase ATP-binding motif

Query Match 9.1%; Score 93.5; DB 2; Length 1734;
Best Local Similarity 26.3%; Pred. No. 15;
Matches 51; Conservative 23; Mismatches 71; Indels 49; Gaps 8;

Qy 25 GLQSYQOAL-----LRISLDKVSRLGRAPSLRRHVLIHTLQQLAALRLAPALPP 79
Db 1308 GSGSFTPLKLSPLGRQLSRPSASBPSPILKR-----VQSAEKLAALAAAE 1357

Qy 80 EPLFLGEEDFSLSATISILRELDTSMDGTEPPQNPVTPGLQNEVP---POPDPVFLEA 136
Db 1358 KKU-APSRKSHLDLPHGELKKEL-----TPRASPLEVVGTRSVLSGKGPLPGKVLQFP 1410

Qy 137 LSSRYLG-----DSGLDDFFLDIDTSAVEKEPARAP-PEPPHN 173
Db 1411 APSRALGTLRODRAERRESLQKQFAIREVDSSEDDTDEEPSOATQEPRLSPHPASHN 1470

Qy 174 LFCAPGSGWENELD 187
Db 1471 LL-PKSGSGEGTEED 1483

RESULT 7
FOLJH2
gag polyprotein - human T-cell lymphotropic virus type 2
N:Alternate names: core polyprotein
C:Contains: core protein p12; core protein p15; core protein p24
C:Species: human T-cell lymphotropic virus type 2, HTLV-2
A>Note: host Homo sapiens (man)
C>Date: 28-Feb-1986 #sequence_revision 28-Feb-1986 #text_change 09-Jul-2004
C:Accession: A03944
R:Shimotohno, K.; Takahashi, Y.; Shimizu, N.; Gojobori, T.; Golde, D.W.; Chen, I.S.Y.; I
Proc. Natl. Acad. Sci. U.S.A. 82, 3101-3105, 1985
A>Title: Complete nucleotide sequence of an infectious clone of human T-cell leukemia vi
A:Reference number: A34042; MUID:85216449; PMID:2582407
A:Accession: A03944
A:Molecule type: DNA
A:Residues: 1-433 <SHI>
A:Cross-references: UNIPROT:P03346; GB:M10060; NID:G329559; PIDN:AAB59884.1; PID:G32956;
A>Note: the authors translated the codon TAC for residues 197 and 249 as Thr
C:Genetics:
A:Gene: gag
C:Superfamily: mammalian retrovirus gag polyprotein II
C:Keywords: core protein; polyprotein
F:1-136/Product: core protein p15 #status predicted <P15>
F:137-214/Product: core protein p24 #status predicted <P24>
F:215-433/Product: core protein p12 #status predicted <P12>

Query Match 9.0%; Score 93; DB 1; Length 433;
Best Local Similarity 23.1%; Pred. No. 2.7;
Matches 40; Conservative 24; Mismatches 67; Indels 42; Gaps 7;

Qy 47 PRAPSLRRHVLIHTLQQLAALRLAPAP-----LPPSPFLGEBDFSLSA 93
Db 13 PKAP---RGLSTHVLNLFQAAAYRLQRPSPDFQQLRFLKALKPTWLNPDYSLLA 69

Qy 94 TI-----GSILRELDTSMDGTEPPQNPVTPLG---LQNEVPPQDPVFLEA-LSSRYLG 143
Db 70 SLIPKPGYGRVVEIINILVKNQVSPSAPADVPVTPICPTTTTPPPPPSPSPHVPVPY-- 127

Qy 144 DSGLDDFFLDIDTSAVEKEPARAPPEPPHNLFCAPGSGWENELDHIMEILIGS 196
Db 128 -----VEPTTQCFFILHPFGAP-----SAHRPWQMKDLQAIKQEVSSS 166

RESULT 8
A82255
hypothetical protein VC0998 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: A82255
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Richardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, I.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20408833; PMID:10952301
A:Accession: A82255
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1621 <HEI>
A:Cross-references: UNIPROT:Q9KTA5; GB:AE004181; GB:AE003852; NID:G9655454; PIDN:AAF9415
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0998

A:Experimental source: fetal kidney; clone DKF2p566N1047
C:Genetics:
A:Note: DKF2p566N1047.1

Query Match 8.5%; Score 88; DB 2; Length 651;
Best Local Similarity 24.4%; Pred. No. 12;
Matches 38; Conservative 26; Mismatches 68; Indels 24; Gaps 7;

QY 27 QSYQALLRLSLDKVRSGLGRAPSLRRHVLHNTLQLOALRLA--PAPALPPEPL-- 82
DB 320 QSQQPPQQQSPQVQVPPQCM-AGPLVTVOSVGLQASSQSVYPAVSPPQHLLP 378
QY 83 -----FLGEEDFSLSATIGLSILRELDTSMDGTGPPQVTVTLGLQNEVPQP--DPVFL 134
DB 379 VSPQTHPMRDD--VATQFGQMTLSROSSGETPEPPSGPVVPSLM-----PQPAQPSYV 432
QY 135 EALSRYLGDGLDDDFLDIDTSVAVEKPARAPPEP 170
DB 433 IASTGQQLPTGTFSG-----SGPTISQCVLQPPESP 463

RESULT 11
T26998
Hypothetical protein Y48B6A.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T26998
R:Mail, M.
Submitted to the EMBL Data Library, September 1999
A:Reference number: Z20297
A:Accession: T26998
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-716 <WIL>
A:Cross-references: UNIPROT:Q9U2A6; EMBL:AL110490; NID:e1542263; PIDN:CAB54442.1; GESP:Y
A:Experimental source: Clone Y48B6A
C:Genetics:
A:Gene: CESP:Y48B6A.6
A:Introns: 38/3; 196/3; 437/2; 460/3; 518/1; 617/3; 673/3

Query Match 8.5%; Score 88; DB 2; Length 716;
Best Local Similarity 25.0%; Pred. No. 14;
Matches 49; Conservative 29; Mismatches 80; Indels 38; Gaps 9;

QY 7 RKHSDLEEEERNEWS--PAGLQSYQALLRLSLDKVQ---RSLGPRAPSLRRHVLHNT 61
DB 219 RSDSEIEEEERKRSEETASFELEAFIMRISKSPVPPVLSIPPPPP---NIPPTI 275
QY 62 LQLOQALLRLAPALPPEPLFLGLEDPSISATIGSI-----LRELD 103
DB 276 PQEVQSPPSPRPTSVPPPIPSGSDVNMDELIESFSDSVFNNSMSPPLPLPRE-- 333
QY 104 TSMOCTE-PPQNVTVTLGLQNEVPQDPVPLEALS--SRYLGDGLDDDFLDIDTSAVEK 161
DB 334 SSLETLEVPEDVETESKV--EASPTPLPKATESINLESSIKALEGLEVKALEAQEASDDR 391
QY 162 EPARAP-----PESP 171
DB 392 PSAPTPIRDSLSLPPPP 407

RESULT 12
A56508
Anucleate primary sterigmata A (apsa) protein - Emericella nidulans
C:Species: Emericella nidulans, Aspergillus nidulans
C:Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 09-Jul-2004
C:Accession: A56508
R:Fischer, R.; Timberlake, W.E.
J. Cell Biol. 128, 485-498, 1995
A:Title: Aspergillus nidulans apsa (anucleate primary sterigmata) encodes a coiled-coil
A:Reference number: A56508; MUID:95164553; PMID:7860626
A:Accession: A56508
A:Status: preliminary

A:Map position: 1

Query Match 8.8%; Score 91.5; DB 2; Length 1621;
Best Local Similarity 27.4%; Pred. No. 20;
Matches 49; Conservative 22; Mismatches 63; Indels 45; Gaps 8;

QY 13 EEEERWESFAGLQSYQALLRLSLDKVRSGLGRAPSLRRHVLHNTLQLOALRLA 72
DB 530 EEDDFDLSCAGVAGDQ-----LDLDFASIEEQ-----DLEQLEA--KAI 570
QY 73 PAPALPPEPLFLGLEDPSISATIGLSILRELDTSMDGT-----PPQNVTVTL- 119
DB 571 DETALDE--ILAEQADPLSESTELDELDDDFDKNEDFPAQTADLLOPEEPILDE 628
QY 120 ----GLQNEVPQDPVPLEA-----LSSRYLGDGLDDDFLDIDTSVAVEKPARAPPE 169
DB 629 EDSQQLNEVLGEFVPELASELIDQNSTELLDELDD--LDLDESIEATEFVSVAPE 685

RESULT 9
AG0365
Probable membrane protein STY0554 [imported] - Salmonella enterica serov
C:Species: Salmonella enterica subsp. enterica serovar typhi
A:Note: This species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AG0365
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
Ch. T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: A50502; MUID:21534947; PMID:11677608
A:Accession: AG0365
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-804 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD04993.1; PID:G16501778; GSPDB:GN00176
C:Genetics:
A:Gene: STY0554
C:Superfamily: Escherichia coli probable membrane protein ybbP

Query Match 8.8%; Score 90.5; DB 2; Length 804;
Best Local Similarity 28.7%; Pred. No. 9.9;
Matches 37; Conservative 16; Mismatches 39; Indels 37; Gaps 7;

QY 38 LDKVRSGLGRAPS-----LRHHVLHNTLQLOALRLAPALPPE 80
DB 478 LDRWQQLPPESPFPYFLINTASEQVAPLKAFIAHQVLPQTFYVIPA-RLTEINGNPT 536
QY 81 PLFLGEEDFSLSATIGLSILRELDTSMDGTGPPQNVTVTLGLQNEVPQDPVFL-BALSS 139
DB 537 ----GQODESLN-----RELNLTWQDTRPAHNPL-----VAGHWPKPGEVSEMERGLAK 581
QY 140 RY----LGDS 145
DB 582 RLNVKLGDS 590

RESULT 10
T42644
Hypothetical protein DKF2p566N1047.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T42644
R:Ottewill, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
Submitted to the Protein Sequence Database, November 1999
A:Reference number: 222231
A:Accession: T42644
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-651 <AAA>
A:Cross-references: UNIPROT:Q9UF93; EMBL:AL133109

A:Molecule type: DNA

A:Residues: 1-1676 <PIS>

A:Cross-references: UNIPROT:Q00083; GB:X62289; NID:G683499; PID:G683500

C:Genetics:

A:Gene: apsa

A:Introns: 149/3

Query Match 8.5%; Score 88; DB 2; Length 1676;
Best Local Similarity 20.5%; Pred. No. 42;
Matches 32; Conservative 19; Mismatches 51; Indels 54; Gaps 6;

QY 36 ISLDKVSRLGSRAPSLRRHVLHNTLQQLQAALRLAPALPPEPLFLGEEDFSLSATI 95

Db 837 VSFEET-PVAPSPPELRTAFFGSGTTEPVAAPVPVPEVALSP-----I 880

QY 96 GSILRELDTSMDGTEPPQNVPTPLGLQNEVPQDPVFL----- 134

Db 881 SS-----QTQTEVIP-----APPPEPIYPPEMAFSQILVEDTLPIAKLP 924

QY 135 EALSSRYLGSGLDDFFLDIDTSAVEKEPARAPPEP 170

Db 925 EPAPERVFAEGTSTDIAVSVSAISSEQTE-PVEP 959

RESULT 13

A34596

C:Species: Homo sapiens (human)

C:Date: 06-Jul-1990 #sequence_revision 09-Oct-1992 #text_change 09-Jul-2004

C:Accession: A34596; S10379

R:Beckmann, H.; Su, L.K.; Radesch, T.

Genes Dev. 4, 167-179, 1990

A:Title: TPE3: a helix-loop-helix protein that activates transcription through the immun

A:Reference number: A34596; MUID:90249724; PMID:2338243

A:Accession: A34596

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-536 <BEC>

A:Cross-references: UNIPROT:PI9532; EMBL:X51330; NID:G37061; PID:G13353

A:Note: the authors translated the codon ACC for residue 433 as Ser, and GAG for residue

C:Genetics:

A:Gene: GDB:TPE3

A:Cross-references: GDB:125870; OMIM:314310

A:Map position: Xp11.23-Xp11.22

C:Keywords: DNA binding; transcription factor

Query Match 8.4%; Score 86.5; DB 2; Length 536;
Best Local Similarity 25.4%; Pred. No. 13;
Matches 47; Conservative 24; Mismatches 55; Indels 59; Gaps 10;

QY 12 LEBEEREWSPAGLQSYQQAALRLSLDKVQSLGPRAPSLRRHVLHNTLQQLQAALRL 71

Db 192 LQEQQR-----SKDLSRQR-----SLEQANRSLQRIQEL-----ELQAIHG 231

QY 72 APAPALPPEPLFLGEEDFSLSATIGSLIRE-LDTSMDGTE-----PPQNVPTPLG 120

Db 232 LPVEGTP-----GLLSLATISDSLKPEQLDIEEGRGARTFHVGGGPAQN----- 279

QY 121 LQNEVPQP--DPVFLEALSSRYLGSGLDDFFLDIDTSAVEKE-----PAR 165

Db 280 APHQPPAPPDALLDLHFFPSDHLGLDG-DFFHLGLEILMEEEGVVGGLSGGALSPLR 338

QY 166 APPEP 170

Db 339 AASDF 343

RESULT 14

T14355

C:Species: Rattus norvegicus (Norway rat)

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C:Accession: T14355

R:Cao, L.; Zhang, L.; Ruiz-Lozano, P.; Yang, Q.; Chien, K.R.; Graham, R.M.; Zhou, M.

J. Biol. Chem. 273, 21077-21083, 1998

A:Title: A novel putative protein-tyrosine phosphatase contains a BRL1-like domain and

A:Reference number: Z18004; MUID:98361981; PMID:9694860

A:Accession: T14355

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-1494 <CAO>

A:Cross-references: UNIPROT:O68902; EMBL:AF077000; NID:G3598973; PID:G3598974; PIDN:AAC

A:Experimental source: brain

C:Genetics:

A:Gene: PTP-TD14

C:Function:

A:Description: may be involved in regulating Ha-ras-dependent cell growth

C:Keywords: phosphoric monoester hydrolase

Query Match 8.4%; Score 86.5; DB 2; Length 1494;
Best Local Similarity 29.1%; Pred. No. 48;
Matches 37; Conservative 19; Mismatches 40; Indels 31; Gaps 9;

QY 62 LQQLQAALRLAPALPPEPLFLGEEDFSLSATIGSLIRE-----LDTSMDGTEPPQNP 115

Db 625 LMQPRAAVPMAPGCVLYPAPVYTSF-----LGLVPSRPQHGVSSPAGVGPQP-P 675

QY 116 VTPLGLQNEVPQP-----PDVFLEALSSRYLGSGLDDFFLDIDT-SAVEKEPARAPPEPP 171

Db 676 I--VGLPSAPPQFSGPE----LAMQVR-PATTTVDSVQAPITSSHMLRPGFAPAPQP- 727

QY 172 HNLFCAP 178

Db 728 ----CFP 730

RESULT 15

B37761

C:Species: Thiobacillus ferrooxidans

C:Date: 31-May-1991 #sequence_revision 31-May-1991 #text_change 09-Jul-2004

C:Accession: B37761

R:Berger, D.K.; Woods, D.R.; Rawlings, D.E.

J. Bacteriol. 172, 4399-4406, 1990

A:Title: Complementatation of Escherichia coli sigma(54) (ntrA)-dependent formate hydrogen

A:Reference number: A37761; MUID:90330545; PMID:2198257

A:Accession: B37761

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-475 <BER>

A:Cross-references: UNIPROT:P24695; GB:M58480; GB:M33831; NID:G154642; PIDN:AAA27379.1;

C:Superfamily: Pseudomonas transcription initiation factor sigma

C:Keywords: DNA binding; transcription regulation

Query Match 8.3%; Score 86; DB 2; Length 475;
Best Local Similarity 21.7%; Pred. No. 12;
Matches 40; Conservative 31; Mismatches 75; Indels 38; Gaps 6;

QY 1 MEGGLKRXHSLDEBEEREWSPAGLQSYQQAALRLSLDKVQR-----SLGPRA--PSLR 53

Db 307 WAGGDAAHKYIQDQLNEARWFIKSLQSRQDTILKVARAIVERQKDFPANGFESMRPMVL 366

QY 54 RHVLHNTLQQLQAALRLAPALPPEPLFLGEEDFSLSATIG-----SILRELDT 104

Db 367 RH--IADAVEMHESVTSRVTNQKYMITPRGLVEFKYFPFSSHVGTDSGGSASATIRALLI 424

QY 105 SMDGTEPPQNVPTPLGLQNEVPQDPVFLEALSSRYLGSGLDDFFLDIDTSAVEKEPA 164

Db 425 RMTQAEADQHPLS-----DAETARVLADOGIQ---IARRTVAKYREAA 464

QY 165 RAPP 168

Db 465 NVPEP 468

Search completed: November 15, 2004, 14:07:21

us-10-069-386-2.rpr

Mon Nov 15 17:33:07 2004

Job time : 40 secs